

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 22:44:12 ; Search time 184 Seconds
(without alignments)
124.172 Million cell updates/sec

Title: US-09-819-144A-2
Perfect score: 268
Sequence: 1 MAAGFRPGACRAGAPTIVL.....SSPARATLGKPLVLDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	268	100.0	52	2 AAY39293	Aay39293 CSA-1 cho
2	77	28.7	209	7 ABM87430	Abm87430 Rice abio
3	74	27.6	16	2 AAY39295	Aay39295 Polypepti
4	73.5	27.4	337	7 ADC32784	Adc32784 Human nov
5	73.5	27.4	675	9 AEA52600	Aea52600 Human Rab
6	73.5	27.4	791	5 ABP65089	Abp65089 Hypoxia-r
7	73.5	27.4	791	8 ADN03640	Adn03640 Antipsori
8	73.5	27.4	791	8 ADP22972	Adp22972 PRO polyp
9	73.5	27.4	863	6 ABP98856	Abp98856 Human str
10	73.5	27.4	863	7 ADC31071	Adc31071 Human nov
11	73.5	27.4	863	7 ADE48296	Ade48296 Human MIC
12	73.5	27.4	863	7 ADE48308	Ade48308 Human MIC
13	73.5	27.4	863	8 ABM82355	Abm82355 Tumour-as
14	72.5	27.1	173	3 ABO76575	Abm76575 Pseudomon
15	71.5	26.7	76	3 AAG28488	Aag28488 Zea mays
16	71	26.5	1938	6 ABP76682	Abp76682 Streptomy
17	69.5	25.9	210	4 AAU48625	Aau48625 Propionib
18	69.5	25.9	210	6 ABM45144	Abm45144 Propionib
19	68	25.4	315	5 ABU05772	Abu05772 M. tuberc
20	68	25.4	315	8 ABM79606	Abm79606 M. tuberc
21	67.5	25.2	135	4 AA001165	Aao01165 Human pol
22	67.5	25.2	137	7 ABO81794	Abm81794 Pseudomon
23	67	25.0	1321	9 ADV97797	Adv97797 Murine pr
24	66	24.6	228	7 ABO79635	Abm79635 Pseudomon

25	65.5	24.4	923	7	ADD18710	Add18710 Human dis
26	65.5	24.4	923	8	ADO19848	Ado19848 Human PRO
27	65.5	24.4	924	8	ADU06294	Adu06294 Novel bro
28	65	24.3	166	7	ABO68030	Abm68030 Pseudomon
29	65	24.3	205	7	ABO82421	Abm82421 Pseudomon
30	65	24.3	345	7	ABO77022	Abm77022 Pseudomon
31	64.5	24.1	146	7	ABO77184	Abm77184 Pseudomon
32	64.5	24.1	149	7	ABO82120	Abm82120 Pseudomon
33	64.5	24.1	309	4	ABG12975	Abg12975 Novel hum
34	64.5	24.1	341	8	ADK71091	Adk71091 Human MP2
35	64	23.9	137	3	RAG28496	Rag28496 Zea mays
36	63.5	23.7	223	8	ADY06811	Ady06811 Plant ful
37	63.5	23.7	377	9	ADY85271	Ady85271 Human ort
38	63.5	23.7	478	7	RAO30817	Rao30817 Human cel
39	63.5	23.7	721	7	ABO79547	Abm79547 Pseudomon
40	63	23.5	53	5	ABP01580	Abp01580 Human ORF
41	63	23.5	128	7	ADM64958	Adm64958 Human pro
42	63	23.5	128	7	ADM03909	Adm03909 Human pro
43	63	23.5	613	7	ABO71209	Abm71209 Pseudomon
44	63	23.5	1129	7	ABO73584	Abm73584 Pseudomon
45	62.5	23.3	85	4	AAU21041	Aau21041 Human nov

ALIGNMENTS

RESULT 1
AAY39293
ID AAY39293 standard; protein; 52 AA.
XX
AC AAY39293;
XX
DT 26-NOV-1999 (first entry)
XX
DE CSA-1 chondrosarcoma associated protein-1.
XX
KW Chondrosarcoma associated protein-1; CSA-1; bone malignancy; diagnosis;
KW Rheumatoid arthritis; inflammatory arthropathy; tumour;
KW cartilage associated polypeptide; CAA-1.
XX
OS Homo sapiens.
XX
PN WO9946382-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US005348.
XX
PR 13-MAR-1998; 98US-00042225.
XX
(RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
XX
Terek RM;
XX
WPI; 1999-551411/46.
XX
N-PSDB; AAZ06793.
XX
New gene encoding a chondrosarcoma associated (CSA) polypeptide, useful for diagnosing bone malignancy.
XX
Claim 10; Page 18; 47pp; English.
XX
This is the human chondrosarcoma associated protein-1 (CSA-1) amino acid sequence. Chondrosarcoma is the second most common form of bone malignancy and occurs in late adulthood and old age. CSA-1 is expressed in a tumour cell line and also in some high grade chondrosarcoma, but not in normal cartilage, or low or intermediate grade tumours. The CSA-1 polynucleotide and polypeptide can be used in methods and compositions for evaluating appropriate treatment and treatment effectiveness of malignancies associated with expression of CSA-1. CSA-1 polynucleotide can be used as a probe to classify cells in terms of their level of CSA-1 expression or as primers for diagnostic PCR analysis in which mutations and allelic variation of CSA-1 can be detected. Transgenic animals

Query Match 28.7%; Score 77; DB 7; Length 209;

RESULT 4

AD	ADC32784	
ID	ADC32784 standard; protein; 337 AA.	
XX		
AC	ADC32784;	
XX		
DT	18-DEC-2003 (first entry)	
XX		
DE	Human novel contig-encoded polypeptide sequence, SEQ ID NO:2866.	
XX		
KW	Human; diagnostic; drug screening; forensics; gene mapping;	
KW	biodiversity assessment; Parkinson's disease; Alzheimer's disease;	
KW	neurodegenerative diseases; anaemia; platelet disorder; wound; burns;	
KW	ulcers; osteoporosis; autoimmune disease; cancer;	
KW	molecular weight marker; food supplement; antiparkinsonian; nootropic;	
KW	neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnery;	
KW	antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;	
KW	gene therapy; chromosome 22q13.1-13.2.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2003029271-A2.	
XX		
PD	10-APR-2003.	
XX		
PF	24-SEP-2002; 2002WO-US030474.	
XX		
PR	24-SEP-2001; 2001US-0324631P.	
XX		
PA	(HYSB-) HYSO INC.	
XX		
PI	Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;	
PI	Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;	
PI	Haley-Vicente D, Drmanac RT;	
XX		
DR	WPI; 2003-371981/35.	
XX		
DR	N-PSDB; ADC32017.	
XX		
PT	New polynucleotide and polypeptide useful for diagnosing, preventing or	
PT	treating conditions such as neurodegenerative diseases, anemias, platelet	
PT	disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or	
PT	cancer.	
XX		
PS	Example 2; SEQ ID NO 2866; 1185pp; English.	
XX		
CC	The invention relates to 971 novel human cDNA sequences (ADC29919-	
CC	ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The	
CC	invention also relates to nucleic acid sequences over 99% identical with	
CC	the novel human cDNAs. The invention additionally encompasses expression	
CC	vectors and host cells comprising a nucleic acid of the invention; the	
CC	recombinant production of a polypeptide of the invention; an antibody	
CC	against a polypeptide of the invention; a method of detecting	
CC	polynucleotides or polypeptides of the invention; and methods of	
CC	identifying a compound which binds to a polypeptide of the invention. The	
CC	invention further discloses methods of preventing, treating or	
CC	ameliorating a medical condition; kits comprising polynucleotide probes	
CC	and/or monoclonal antibodies for carrying out the methods of the	
CC	invention; methods for the identification of compounds that modulate the	
CC	expression or activity of the polynucleotide and/or polypeptide; and 767	
CC	contlig sequences corresponding to the cDNA sequences of the invention	
CC	(ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC3262	
CC	-ADC33394). The nucleic acids and polypeptides of the invention are	
CC	useful in diagnostics, drug screening, forensics, gene mapping, in the	
CC	identification of mutations responsible for genetic disorders or other	
CC	traits, for assessing biodiversity, and in producing many other types of	
CC	data and products dependent on DNA and amino acid sequences. They are	
CC	also used for treating diseases such as Parkinson's disease, Alzheimer's	
CC	disease and other neurodegenerative diseases, anaemia, platelet	
CC	disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or	
CC	cancer. The nucleic acids may also be used as hybridisation probes or	
CC	primers, and in the recombinant production of a protein. The polypeptide	
CC	are also useful in generating antibodies, as molecular weight markers,	
CC	and as food supplements. The present sequence represents a human contig-	
CC	encoded polypeptide sequence used in an example of the invention. Note:	

```
ABP65089
ID ABP65089 standard; protein; 791 AA.
XX AC
XX AC ABP65089;
XX DT
XX DT 12-NOV-2002 (first entry)
XX DE
XX DE Hypoxia-repressed protein #12.
XX KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
XX KW antiinflammatory; vulnary; gynecological; ophthalmological; vaccine;
XX KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
XX KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
XX KW precleptemia; atherosclerosis; inflammatory condition; wound healing;
XX KW inflammation; erythropoiesis; hair loss; human.
XX OS
XX OS Homo sapiens.
XX PN WO200246465-A2.
XX PD
XX PD 13-JUN-2002.
XX PF
XX PF 10-DEC-2001; 2001WO-GB005458.
XX PR
XX PR 08-DEC-2000; 2000GB-00030076.
XX PR 08-FEB-2001; 2001GB-00003156.
XX PR 25-OCT-2001; 2001GB-00025666.
XX PA
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX FI
XX FI White J, Mundy CR, Ward NR, Krige D, Kingeman SM, Harris RA;
XX FI Rayner WN;
XX DR
XX DR WPI; 2002-627238/67.
XX PT
XX PT Identifying a gene involved in disease for treating hypoxia-regulated
XX PT conditions, comprises comparing the transcriptome/proteome of a
XX PT types under different conditions and identifying a differentially
XX PT regulated gene.
XX PS
XX PS Claim 13; Page 290; 538pp; English.
XX CC
XX CC The present invention relates to methods for identifying genes and
XX CC proteins that are implicated in a specific disease or physiological
XX CC condition. The method comprises comparing the transcriptome/proteome of a
XX CC specialised cell type implicated in a disease or condition with that of a
XX CC second specialised cell type, under two experimental conditions, and
XX CC identifying a gene that is differentially regulated in the two
XX CC specialised cell types under experimental conditions. ABV77873-ABV78116
XX CC and ABP65061-ABP65257 were identified using the methods of the invention.
XX CC The coding sequences and proteins are useful for treating a disease in a
XX CC patient, for manufacture of a medicament for treating hypoxia-regulated
XX CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
XX CC biological response to hypoxia conditions, or hypoxic-associated
XX CC pathology in a patient. The coding sequences and proteins are also useful
XX CC for monitoring the therapeutic treatment of a disease or physiological
XX CC condition, such as cancer, ischaemic conditions, reperfusion injury,
XX CC retinopathy, neonatal stress, precleptemia, atherosclerosis, inflammatory
XX CC conditions, wound healing, inflammation, erythropoiesis or hair loss
XX SQ
XX SQ Sequence 791 AA;
XX Query Match 27.4%; Score 73.5; DB 5; Length 791;
XX Best Local Similarity 45.7%; Pred. No. 1.9;
XX Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;
XX QY
XX QY 2 AAGPRGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
XX Db
XX Db 516 SSGFQPAKPCSGATPTPLLAVGDRSPVSPGSSSP 550
XX RESULT 7
XX ADN03640
XX Query Match 27.4%; Score 73.5; DB 8; Length 791;
XX Best Local Similarity 45.7%; Pred. No. 1.9;
XX Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;
XX QY
XX QY 2 AAGPRGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
XX Db
XX Db 516 SSGFQPAKPCSGATPTPLLAVGDRSPVSPGSSSP 550
XX RESULT 8
XX ADP22972
XX ADP22972 ADP22972 standard; protein; 791 AA.
XX AC
XX AC ADP22972;
XX DT
XX DT 18-NOV-2004 (first entry)
XX DE
XX DE PRO polypeptide SEQ ID NO:66.
XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
XX KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
XX KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX OS
XX OS Unidentified.
XX PN WO2004041170-A2.
XX PD
XX PD 21-MAY-2004.
XX PF
XX PF 30-OCT-2003; 2003WO-US034312.
XX XX
```


KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 22q13.1-13.2.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 XX
 XX 10-APR-2003.
 PD
 XX 24-SEP-2002; 2002WO-US030474.
 PF
 XX 24-SEP-2001; 2001US-0324631P.
 PR
 XX (HYSE-) HYSEQ INC.
 FA
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 DR WPI; 2003-371981/35.
 DR N-PSDB; ADC30100.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Claim 20; SEQ ID NO 1153; 1185pp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30899) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 863 AA;

Query Match 27.4%; Score 73.5; DB 7; Length 863;
 Best Local Similarity 45.7%; Pred. No. 2.1; Indels 1; Gaps 1;
 Matches 16; Conservative 5; Mismatches 13;

Qy 2 AAGPRGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35
 Db 588 SSGPQPAKPCSGATPTPLLLVGRSPVPSPGSSSP 622

RESULT 11

ADE48296
 ID ADE48296 standard; protein; 863 AA.

XX AC ADE48296;

XX DT 29-JAN-2004 (first entry)

XX DE Human MICAL-1.

KW MICAL domain; calponin homology domain; LIM domain; plexin;
 KW Neuroprotective; Analgesic; Antiparkinsonian; vulnerary; Tranquillizer;
 KW Vasotropic; Cardiant; Nootropic; Gene therapy; Antisense gene therapy;
 KW Cell therapy; Monoxygenase inhibitor; neurological condition.

XX OS Homo sapiens.

XX PN WO2003066821-A2.

XX PD 14-AUG-2003.

XX PF 04-FEB-2003; 2003WO-US003551.

XX PR 04-FEB-2002; 2002US-0354178P.

XX PR 30-MAY-2002; 2002US-0384302P.

XX PR 13-JUN-2002; 2002US-0388325P.

XX PA (UWJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

XX PI Kolodkin AL, Terman JR, Mao T, Pasterkamp RJ, Yu H;

XX DR WPI; 2003-731494/69.

XX DR N-PSDB; ADE48295.

XX PT New polypeptide comprising an N-terminal MICAL domain, a calponin

XX PS Claim 4; SEQ ID NO 2; 207pp; English.

CC The present sequence relates to an isolated polypeptide comprising an N-
 CC terminal MICAL domain, a calponin homology domain, a LIM domain, a
 CC proline rich region or a plexin interacting region, where the polypeptide
 CC has monooxygenase activity. The polypeptide is useful for treating a
 CC neurological condition, e.g., spinal cord injury, traumatic brain injury,
 CC neuropathic pain, Parkinson's disease, Amyotrophic Lateral Sclerosis,
 CC ischemic injury, Alzheimer's disease, multiple sclerosis or neuropathy
 CC resulting from a stroke. The present sequence represents a MICAL protein.

XX SQ Sequence 863 AA;

Query Match 27.4%; Score 73.5; DB 7; Length 863;

Best Local Similarity 45.7%; Pred. No. 2.1; Indels 1; Gaps 1;

Matches 16; Conservative 5; Mismatches 13;

Qy 2 AAGPRGAPCRAGAPTIVLTSGRRQTL-SHGSSSP 35

Db 588 SSGPQPAKPCSGATPTPLLLVGRSPVPSPGSSSP 622

RESULT 12

ADE48308
 ID ADE48308 standard; protein; 863 AA.

XX AC ADE48308;

XX DT 29-JAN-2004 (first entry)

XX	Human MICAL-1 like-protein.
DE	MICAL domain; calponin homology domain; LIM domain; plexin;
XX	Neuroprotective; Analgesic; Antiparkinsonian; Vulnery; Tranquillizer;
KW	Vasotropic; Cardiant; Nootropic; Gene therapy; Antisense gene therapy;
KW	Cell therapy; Monoxygenase inhibitor; neurological condition.
XX	
OS	Homo sapiens.
XX	
PN	WC02003066821-A2.
XX	
PD	14-AUG-2003.
XX	
PF	04-FEB-2003; 2003WO-US003551.
XX	
PR	04-FEB-2002; 2002US-0354178P.
XX	
PP	30-MAY-2002; 2002US-0384302P.
PR	13-JUN-2002; 2002US-0388325P.
XX	
PA	(UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX	
PI	Kolodkin AL, Terman JR, Mao T, Pasterkamp RJ, Yu H;
XX	
PT	WFI; 2003-731494/69.
XX	N-PSDE; ADE48307.
XX	
PT	New polypeptide comprising an N-terminal MICAL domain, a calponin
PT	homology domain, a LIM domain, a proline rich region or a plexin
PT	interacting region, useful for treating a neurological condition, e.g.,
PT	Alzheimer's disease.
XX	
XX	Claim 72; SEQ ID NO 14; 207pp; English.
PS	
CC	The present sequence relates to an isolated polypeptide comprising an N-
CC	terminal MICAL domain, a calponin homology domain, a LIM domain, a
CC	proline rich region or a plexin interacting region, where the polypeptide
CC	has monoxygenase activity. The polypeptide is useful for treating a
CC	neurological condition, e.g., spinal cord injury, traumatic brain injury,
CC	neuropathic pain, Parkinson's disease, Amyotrophic Lateral Sclerosis,
CC	ischemic injury, Alzheimer's disease, multiple sclerosis or neuropathy
CC	resulting from a stroke. The present sequence represents a MICAL protein.
XX	
SQ	Sequence 863 AA;
	Query Match 27.4%; Score 73.5; DB 7; Length 863;
	Best Local Similarity 45.7%; Pred. No. 2.1;
	Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;
Qy	2 AAGPRGAPCRAGAPTIVLTSGRQTLSHGSSSP 35
	: : : : : :
Dd	588 SSGQPAPKPCSGATPTLLVGDSPVPSPGSSSP 622
RESULT 13	
ABM82355	ID ABM82355 standard; protein; 863 AA.
XX	
AC	ABM82355;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Tumour-associated antigenic target (TAT) polypeptide PRO83279, SEQ:6048.
XX	
KW	Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW	tumour; diagnosis; cell proliferative disorder; breast cancer;
KW	colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW	central nervous system cancer; bladder cancer; pancreatic cancer;
KW	cervical cancer; melanoma; leukaemia; hybridisation probe;
KW	chromosome identification; chromosome mapping; gene mapping;
KW	gene therapy; cytostatic.
XX	
OS	Homo sapiens.

XX The invention relates to avilamycin derivatives (I) with antibacterial,
 CC virucide, protozoacide and fungicide activity. (I) are useful for
 CC treatment of infections (bacterial, viral, protozoal or fungal), in human
 CC or veterinary medicine, particularly where caused by *Staphylococcus*
 CC aureus. (I) are more hydrophilic than known avilamycins. The present
 CC sequence is that of an avilamycin synthesis enzyme from the *Streptomyces*
 CC viridochromogenes Avilamycin A biosynthetic gene cluster (AB237515-
 CC AB237516)
 CC
 XX SQ Sequence 19938 AA;
 Query Match 26.5%; Score 71; DB 6; Length 1938;
 Best Local Similarity 39.6%; Pred. No. 2e+02;
 Matches 19; Conservative 2; Mismatches 15; Indels 12; Gaps 2;
 QY 2 AAGPRPGACRCAGATTIVLTSGRQTLSHGS-----SSPARATLGKP 43
 Db 3679 ARPPRGRHCKGGAP-----GKSGRSRGTAALAUSSPGRTARS GP 3720
 RESULT 17
 AAU48625
 ID AAU48625 standard; protein; 210 AA.
 AC AAU48625;
 XX
 XX 27-FEB-2002 (first entry)
 XX
 XX Propionibacterium acnes immunogenic protein #9521.
 XX
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 XX Propionibacterium acnes.
 XX
 XX WO200181581-A2.
 XX
 XX 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012865.
 XX
 XX 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
 PI
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 XX Example 1; SEQ ID NO 9820; 1069pp; English.
 PS
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX SQ Sequence 210 AA;
 Query Match 25.9%; Score 69.5; DB 4; Length 210;
 Best Local Similarity 38.3%; Pred. No. 1.3;
 Matches 18; Conservative 5; Mismatches 19; Indels 5; Gaps 2;
 QY 2 AAGPRPGACRCAGATTIVL---TSGRQTLSHGSSSPARA-TLGKP 43
 Db 85 AVHPRHCGCTRRGRSVVAVSRSDRGSRTPAPGHPRPARSRRLGVP 131
 RESULT 18
 ABM45144
 ID ABM45144 standard; protein; 210 AA.
 XX
 XX ABM45144;
 AC
 XX 20-OCT-2003 (first entry)
 XX
 XX Propionibacterium acnes predicted ORF-encoded polypeptide #9820.
 XX
 XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 KW
 XX Propionibacterium acnes.
 OS
 XX WO2003033515-A1.
 XX
 XX 24-APR-2003.
 XX
 XX 11-OCT-2002; 2002WO-US032727.
 XX
 XX 15-OCT-2001; 2001US-00978825.
 PR
 XX (CORI-) CORIXA CORP.
 PA
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallie-Douglasse J;
 PI
 XX WPI: 2003-381789/36.
 DR N-PSDB; ACF64472.
 DR
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 XX Example 1; SEQ ID NO 9820; 1481pp; English.
 PS
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a

CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridization. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 210 AA;

Query Match 25.9%; Score 69.5; DB 6; Length 210;
 Best Local Similarity 38.3%; Pred. No. 1.3;
 Matches 18; Conservative 5; Mismatches 19; Indels 5; Gaps 2;

QY 2 AAGPRGAPCRAGAPTIVL---TSGRRQTLGSHGSSPARA-TLGKP 43
 DB 85 AVHPRHGCTRGSRVAVVSRSDSGRSNTAPHGHPARPSRLGVP 131

RESULT 19
 ABU05772
 ID ABU05772 standard; protein; 315 AA.

AC ABU05772;
 XX
 DT 08-APR-2003 (first entry)
 XX
 DE M. tuberculosis and M. leprae marker protein #423.
 XX
 KW Mycobacterioses; survival; virulence; protective antigen; vaccine;
 KW mycobacterial disease; tuberculosis; leprosy.

OS Mycobacterium tuberculosis.
 OS Mycobacterium leprae.

PN WO200274903-A2.

XX 26-SEP-2002.

XX 22-FEB-2002; 2002WO-IB001973.

PR 22-FEB-2001; 2001US-0270123P.

XX (INSP) INST PASTEUR.

PA Cole S;

PI WPI; 2002-759885/82.

XX Identifying and selecting genes for survival or virulence of mycobacteria
 PT by a comparative genomic analysis of the sequences of Mycobacterium
 PT tuberculosis and M. leprae.

PS Claim 17; Page 623-624; 874pp; English.

XX This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC leprae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a marker protein from

CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
 CC method of the invention
 XX
 SQ Sequence 315 AA;

Query Match 25.4%; Score 68; DB 5; Length 315;
 Best Local Similarity 28.2%; Pred. No. 3.4;
 Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;
 QY 1 MAAGPRGAPC-----RAGAPT-----IVLTSGRRQTL 28
 DB 51 LVAGSRPGPIIGVNLNLPSPRGAGGAAELVVHPQSRRRGIGTAMARAALAKTGRNQFW 110
 QY 29 SHGSSSPARAT 39
 DB 111 AHGTLDPARAT 121

RESULT 20
 ABM79606
 ID ABM79606 standard; protein; 315 AA.

XX AC ABM79606;

XX 22-APR-2004 (first entry)

XX M tuberculosis MshD protein.

XX Enzyme; mycothiol synthesis; MshC; MshD; MshA; infection;
 KW cysteine:glucosaminyl inositol ligase; antibacterial;
 KW acetyl-CoA:Cys-GlcN-Ins acetyltransferase.

OS Mycobacterium tuberculosis.

XX WO2003089585-A2.

XX 30-OCT-2003.

PF 15-APR-2003; 2003WO-US011539.

PR 15-APR-2002; 2002US-0373079P.

XX 19-APR-2002; 2002US-0373890P.

XX (REGC) UNIV CALIFORNIA.

PA (UYBR-) UNIV BRITISH COLUMBIA.

PI Sareen D, Newton GL, Fahey RC, Buchmeier N, Steffek M, Av-Gay Y;

XX Rawat M, Koledin T;

XX WPI; 2004-042359/04.

XX Identifying inhibitors of mycothiol biosynthesis, useful as antibacterial

PT agents and for increasing sensitivity to antibiotics, also mycothiol-

PT defective mutant actinomycetes.

PS Claim 38; Fig 10A; Opp; English.

XX The present invention relates to a method for identifying an inhibitor of
 CC cysteine:glucosaminyl inositol ligase (CGIL), which comprises treating
 CC test compound with such a ligase in presence of a cysteine and a
 CC glucosaminyl inositol, or their derivatives and determining if ligation
 CC of cysteine and glucosaminyl inositol has occurred. Agents which inhibit
 CC cysteine:glucosaminyl inositol ligase, or other enzymes involved in
 CC biosynthesis of mycothiol (MSH), are used to reduce virulence, increase
 CC antibiotic sensitivity and inhibit growth of MSH-producing bacteria,
 CC especially for treatment of infections by Mycobacterium species, but also
 CC e.g. Corynebacterium diphtheriae or Actinomyces israelii. New mutant
 CC forms of actinomycetes that have mutated MSH-synthesis genes and can only
 CC survive briefly in white blood cells are useful as vaccines. The present
 CC sequence is a polypeptide which is/forms part of an enzyme relevant to
 CC the invention

XX Sequence 315 AA;

```

Query Match      25.4%; Score 68; DB 8; Length 315;
Best Local Similarity 28.2%; Pred. No. 3.4;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;

Qy 1 MAAGPRPGAPC-----RAGAPT-----IVLTSGRRQTL 28
Db 51 LVAGSRPGPIIGVNLSPPRGAGGAAELVHPQSRRRGIGTAMARAALAKTAGRNQFW 110
Qy 29 SHGSSSPARAT 39
Db 111 AHGTLDPARAT 121

RESULT 21
AAO01165
ID AAO01165 standard; protein; 135 AA.
XX
AC AAO01165;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 15057.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX
PS 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-514838/56.
XX
N-PSDB; AA181096.

Isolated nucleic acids and polypeptides, useful for preventing diagnosing
and treating e.g. leukemia, inflammation and immune disorders.

Claim 20; SEQ ID NO 15057; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and
the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 135 AA;

Query Match      25.2%; Score 67.5; DB 4; Length 135;
Best Local Similarity 43.2%; Pred. No. 1.4;
Matches 19; Conservative 4; Mismatches 12; Indels 9; Gaps 2;

Qy 2 AAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLKGKPLV 45

```

```

Db 92 SALPAPGAPRRHGHPLMVL-----QTLPGHPHTPAR-----KPYI 136

RESULT 22
ABO81794
ID ABO81794 standard; protein; 137 AA.
XX
AC ABO81794;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #13969.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX
PS 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
WPI; 2003-615309/58.
XX
N-PSDB; ABD15365.

Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
useful as molecular targets for diagnostics, prophylaxis and treatment of
pathological conditions resulting from bacterial infection.

Disclosure; SEQ ID NO 30540; 455pp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the
polynucleotides encoding them. The sequences are useful in diagnosis and
therapy of pathological conditions, as molecular targets for diagnostics,
prophylaxis and treatment of pathological conditions resulting from a
bacterial infection, for evaluating a compound, such as a polypeptide,
for the ability to bind a P. aeruginosa nucleic acid, as components of
effective antibacterial targets, as targets for antibacterial drugs,
including anti-P. aeruginosa drugs, as templates for recombinant
production of P. aeruginosa-derived peptides or polypeptides, as target
components for diagnosis and/or treatment of P. aeruginosa-caused
infection, and in detection of P. aeruginosa sequences or other sequences
of Pseudomonas species using biochip technology. Sequences ABO67826-
ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
sequence data for this patent did not form part of the printed
specification but was obtained in electronic format from USPTO at
seqdata.uspto.gov/sequence.html

XX
SQ Sequence 137 AA;

Query Match      25.2%; Score 67.5; DB 7; Length 137;
Best Local Similarity 36.0%; Pred. No. 1.4;
Matches 18; Conservative 5; Mismatches 12; Indels 15; Gaps 2;

Qy 2 AAGPRPG-----APCRAGAPTIVLTSGRRQTLSHGSSSPAR 37
Db 44 APAPRPGRGPSIDRRGSPGRPCLAG-PSAASPCGPRSRISAGSRSPAR 92

RESULT 23
ADV97797
ID ADV97797 standard; protein; 1321 AA.
XX
AC ADV97797;

```


DT 15-JAN-2004 (first entry)
 XX Human disease related protein SeqID141.
 DE human; disease state; cytostatic; antiinflammatory; ophthalmological;
 XX antiarteriosclerotic; vulnerary; gene therapy;
 KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
 KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
 KW glucose transportation; catecholamine synthesis; iron transport;
 KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
 KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
 KW inflammatory condition; wound healing.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO2003018621-A2.
 XX 27-MAY-2004.
 XX 06-MAR-2003.
 XX 23-AUG-2002; 2002WO-GB003892.
 XX 23-AUG-2001; 2001GB-00020558.
 XX 05-OCT-2001; 2001GB-00024037.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
 PI WPI; 2003-290046/28.
 DR N-PSDB; ADD18711.
 XX New substantially purified polypeptide, useful for diagnosing or treating
 PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
 PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
 PT wound healing.
 XX Claim 25; SEQ ID NO 141; 424pp; English.
 PS This invention relates to novel human genes and gene product which are
 CC implicated in certain disease states. Compounds which modulate the
 CC proteins of the invention may have cytostatic, antiinflammatory, the
 CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
 CC sequences of the invention may be useful for gene therapy. The invention
 CC may be useful for diagnosing or treating a hypoxia-regulated condition,
 CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,
 CC erythropoiesis, or the biological response to hypoxia conditions
 CC including processes such as glycolysis, gluconeogenesis, glucose
 CC transportation, catecholamine synthesis, iron transport or nitric oxide
 CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
 CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
 CC inflammatory conditions or wound healing. The present sequence is that of
 CC a disease related protein of the invention.
 XX Sequence 923 AA;
 SQ Query Match 24.4%; Score 65.5; DB 7; Length 923;
 Best Local Similarity 40.0%; Pred. No. 26;
 Matches 14; Conservative 5; Mismatches 7; Indels 9; Gaps 1;
 QY 3 AGPRPGAPCAGAPTIVLTSGRRQTLHGSSSPAR 37
 Db 45 AGPARACPCRAGTPTS-----TWSSPASTPSR 70
 RESULT 26
 ADO19848
 ID ADO19848 standard; protein; 923 AA.
 AC ADO19848;
 XX 12-AUG-2004 (first entry)
 DT Human PRO polypeptide #386.
 DE

XX Human; PRO; immune related disorder; systemic lupus erythematosus;
 KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
 KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
 KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
 KW diabetes mellitus; renal disease; demyelinating disease;
 KW central nervous system; peripheral nervous system;
 KW demyelinating polyneuropathy; Guillain-Barre syndrome;
 KW chronic inflammatory demyelinating polyneuropathy.
 XX Homo sapiens.
 XX WO2004043361-A2.
 XX 27-MAY-2004.
 XX 06-NOV-2003; 2003WO-US035268.
 XX 08-NOV-2002; 2002US-0425235P.
 XX (GETH) GENENTECH INC.
 XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 XX WPI; 2004-420067/39.
 DR N-PSDB; ADO19847.
 XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
 PT treating an immune related disorder such as systemic lupus erythematosus,
 PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
 PT spondyloarthritis.
 XX Claim 7; SEQ ID NO 772; 1731pp; English.
 XX The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the
 CC central or peripheral nervous system, demyelinating polyneuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polyneuropathy. This sequence represents a human PRO polypeptide of the
 CC invention.
 XX Sequence 923 AA;
 SQ Query Match 24.4%; Score 65.5; DB 8; Length 923;
 Best Local Similarity 40.0%; Pred. No. 26;
 Matches 14; Conservative 5; Mismatches 7; Indels 9; Gaps 1;
 QY 3 AGPRPGAPCAGAPTIVLTSGRRQTLHGSSSPAR 37
 Db 45 AGPARACPCRAGTPTS-----TWSSPASTPSR 70
 RESULT 27
 ADO06294
 ID ADO06294 standard; protein; 924 AA.
 XX ADO06294;
 XX 27-JAN-2005 (first entry)
 DT Novel bronchial cancer-associated human protein SeqID518.
 DE bronchial cancer; cytostatic; tumour-associated protein;
 KW cancer detection; metastasis; tumour; human.
 XX Homo sapiens.
 OS

Query match 24.1%; Score 84.5; DB 7; Length 145;
Best Local Similarity 41.7%; Pred. No. 3.9;

50	sequence 303 aa,	Query Match	24.1%;	Score 64.5;	DB 4;	Length 309;
		Best Local Similarity	34.0%;	Pred. No. 9.5;		

Qy 5 PRPGAP-----CRAGAPT-----IVLTSGRRQTLTSHGSSSPARATLGKPLV 45
Db 101 PRPSAPPPRPSRGSSEAEELSLSARTKTRSYGSTASVRAPLGAGVI 150
RESULT 34
ADK71091
ID ADK71091 standard; protein; 341 AA.
XX AC ADK71091;
XX DT 06-MAY-2004 (first entry)
XX DE Human MP21 polypeptide.
XX MP21; p21; p21 modulator; cytostatic; adenylate cyclase modulator;
XX antisense therapy; cancer; human.
XX OS Homo sapiens.
XX WO2004015073-A2.
XX 19-FEB-2004.
XX 06-AUG-2003; 2003WO-US024575.
XX 07-AUG-2002; 2002US-0401739P.
XX (EXEL-) EXELIXIS INC.
XX Francis-Lang H, Friedman L, Kidd T, Roche S, Adamkewicz JI;
XX Platt DM;
XX WPI; 2004-180657/17.
XX N-PSDB; ADK71085.
XX Identifying a candidate p21 pathway modulating agent using an assay
XX system having a modulator of p21 (MP21) polypeptide or nucleic acid,
XX useful for diagnosing or treating cancer, such as colon, kidney, lung and
XX breast cancer.
XX Example 1; SEQ ID NO 8; 97pp; English.
XX The invention relates to identifying a candidate p21 pathway modulating
XX agent. The method involves providing an assay system comprising a
XX modulator of p21 (MP21) polypeptide or nucleic acid, contacting the assay
XX system with a test agent, where in its presence the system provides a
XX reference activity. The methods and compositions of the present invention
XX are useful for the diagnosis and/or treatment of diseases or conditions
XX associated with aberrant expression or activity of the p21 pathway, such
XX as cancer, preferably colon or head and neck cancer. The present sequence
XX represents a human MP21 polypeptide.
XX Sequence 341 AA;
Query Match 24.1%; Score 64.5; DB 8; Length 341;
Best Local Similarity 34.0%; Pred. No. 11;
Matches 17; Conservative 11; Mismatches 13; Indels 9; Gaps 2;
Qy 5 PRPGAP-----CRAGAPT-----IVLTSGRRQTLTSHGSSSPARATLGKPLV 45
Db 134 PRPSAPPPRPSRGSSEAEELSLSARTKTRSYGSTASVRAPLGAGVI 183
RESULT 35
AAG28496
ID AAG28496 standard; protein; 137 AA.
XX AC AAG28496;
XX DT 17-OCT-2000 (first entry)
XX DE Zea mays protein fragment SEQ ID NO: 33735.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; Corn.
XX Zea mays subsp. mays.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 18-MAY-1999; 99US-0134370P.
XX 19-MAY-1999; 99US-0134768P.
XX 20-MAY-1999; 99US-0134941P.
XX 21-MAY-1999; 99US-0135124P.
XX 24-MAY-1999; 99US-0135353P.
XX 25-MAY-1999; 99US-0135629P.
XX 27-MAY-1999; 99US-0136021P.
XX 28-MAY-1999; 99US-0136392P.
XX 01-JUN-1999; 99US-0136782P.
XX 03-JUN-1999; 99US-0137222P.
XX 04-JUN-1999; 99US-0137528P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 21-JUN-1999; 99US-0139763P.
XX 22-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.

```
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.

PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 23.9%; Score 64; DB 3; Length 137;
Best Local Similarity 42.2%; Pred. No. 4.2;
Matches 19; Conservative 4; Mismatches 16; Indels 6; Gaps 3;

QY 5 PRPGAPCRAGAPTIVLTSGRQTLSHGSSSPARATLKGKPLVLDY 49
   |||||
Db 26 PRQGSFCRTRST---TSNRPTL--GRPRPTRSLA-PVRKNGF 64

RESULT 36
ADY06811
ID ADY06811 standard; protein; 223 AA.
XX ADY06811;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 62626.
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
```

KW yield; plant growth; plant development; seed oil; protein yield;
 XX protein content.
 OS Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LTUJ/) LTU J.

XX (ZHOU/) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABA/) TABASKA J E.

XX (CAOY/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to
 XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 XX pests, for conferring increased resistance to plant disease, or for
 XX improving yield.

XX Claim 1; SEQ ID NO 62626; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 XX polynucleotide consisting of a sequence encoding an amino acid sequence
 XX available in electronic form from the US patent office at
 XX ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
 XX of the invention are also useful in physical arrays of molecules and as
 XX plant breeding markers. The recombinant DNA construct is useful for
 XX improving plant tolerance to cold, heat, drought, herbicides, extreme
 XX osmotic conditions, pathogens or pests, for manipulating growth rate in
 XX plant cells by modification of the cell cycle pathway, for conferring
 XX increased resistance to plant disease, for producing galactomannan,
 XX lignin or plant growth regulators, for increasing the rate of homologous
 XX recombination in plants, for improving yield by modification of
 XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 XX or by providing improved plant growth and development under at least one
 XX stress condition or for modifying seed oil or protein yield and/or
 XX content. This is the amino acid sequence of a plant full length insert
 XX polypeptide that can be used in the recombinant DNA construct of the
 XX invention.

XX Sequence 223 AA;

Query Match 23.7%; Score 63.5; DB 8; Length 223;
 Best Local Similarity 34.8%; Pred. No. 8.7;
 Matches 16; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

QY 2 AAGPRPGAPCR-AGAPTIVLTSGRRQTLTSHGSSSPARATLKGKPLVL 46

DB 106 SAGPRPSTSRPPWAAAAAARARRRLLPRGASPRPRRPRRL 151

RESULT 37

ADY85271

ID ADY85271 standard; protein; 377 AA.

XX AC ADY85271;

XX 02-JUN-2005 (first entry)

XX Human orthologue of CD4.4 gene-encoded protein.

XX Cell cycle; restenosis; vasotropic; cardiovascular disease; neoplasm;

KW lymphoma; lung tumor; colon tumor; ovary tumor; cytostatic;
 KW RNA interference; gene silencing; antibody therapy; drug screening;
 KW CD4.4.

OS Homo sapiens.

XX WO2005026359-A2.

XX 24-MAR-2005.

XX 15-SEP-2004; 2004WO-EP010307.

XX 15-SEP-2003; 2003US-0502633P.

XX (CENI-) CENTX BIOSCIENCE GMBH.

XX Echeverri C, Hyman A, Goenczy P, Soennichsen B, Jones S, Walsh A;

XX Kobki L;

XX WPI; 2005-233496/24.

XX N-PSDB; ADY85270.

XX GENBANK; CAD38936.

XX Use of a nucleic acid molecule, polypeptide, or an antibody directed
 XX against the polypeptide, for manufacturing a medicament that inhibits or
 XX activates spindle formation or microtubule function during cell division.
 XX Claim 8; SEQ ID NO 36; 184pp; English.

XX The invention relates to the functional role of several Caenorhabditis
 XX elegans genes and of their corresponding gene products in spindle
 XX formation or microtubule formation during cell division that can be
 XX identified by means of RNA interference (RNAi), and to the identification
 XX and isolation of functional orthologues of these genes and their
 XX biologically functional derivatives. The invention also relates to the
 XX use of these genes and gene products, including the orthologues, in the
 XX development or isolation of anti-proliferative agents, particularly their
 XX use in screening assays, and their use for diagnosis and treatment of
 XX proliferative and other diseases. In particular, the invention relates to
 XX the use of small interfering RNAs (siRNAs) derived from the genes for the
 XX treatment of proliferative diseases. The proliferative disease is
 XX especially coronary restenosis or a neoplastic disease selected from
 XX lymphoma, lung cancer, colon cancer, ovarian cancer and breast cancer
 XX (all claimed). The present sequence is that of the human orthologue of
 XX the C. elegans CD4.4 protein ADY85267. Characterization of the nematode
 XX gene by RNAi indicated a requirement for CD4.4 gene function in spindle
 XX formation or microtubule formation during cell division. Based on the
 XX extremely high sequence conservation at the protein level, these
 XX functions are probably also found in humans. The use of a polypeptide
 XX having the present sequence for the manufacture of a medicament for the
 XX inhibition of spindle formation or microtubule formation during cell
 XX division is claimed.

XX Sequence 377 AA;

Query Match 23.7%; Score 63.5; DB 9; Length 377;

Best Local Similarity 40.5%; Pred. No. 16;

Matches 17; Conservative 2; Mismatches 20; Indels 3; Gaps 1;

QY 5 PR---PCAPCRAGAPTIVLTSGRRQTLTSHGSSSPARATLKGK 43

DB 290 PRFGYPTMGASGPGYPLRGRAPSPFGYQQSPYPATGGKP 331

RESULT 38

AAO30817

ID AAO30817 standard; protein; 478 AA.

XX AC AAO30817;

XX 22-SEP-2003 (first entry)

XX Human cell adhesion and extracellular matrix protein (CADECM)-7.

XX Human; cell adhesion and extracellular matrix protein; immune disorder;
KW CADECM; cancer; gene therapy.
XX Homo sapiens.
XX WO2003047526-A2.
XX PD 12-JUN-2003.
XX PF 26-NOV-2002; 2002WO-US038437.
XX PR 30-NOV-2001; 2001US-0334343P.
XX PR 07-DEC-2001; 2001US-0340278P.
XX PR 04-JAN-2002; 2002US-0345069P.
XX PR 25-JAN-2002; 2002US-0351352P.
XX PR 14-FEB-2002; 2002US-0357168P.
XX PR 29-MAR-2002; 2002US-0369128P.
XX PR 05-APR-2002; 2002US-0370802P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Baughn MR, Becha SD, Bhatia U, Blake JJ, Borowsky ML, Burrill JD;
PI Delegeane AM, Elliott VS, Gandhi AR, Gietzen KJ, Gorvad AE;
PI Griffin JA, Ho A, Jin P, Kable AE, Lal PG, Lee EA, Lee S, Lee SY;
PI Marquis JP, Lehr-Mason PM, Ramkumar J, Richardson TW, Sprague WW;
PI Swatnkar A, Tang TV, Tran B, Tran UK, Chawla NK, Warren BA, Xu Y;
PI Yue H, Zheng W;
XX WPI: 2003-513695/48.
XX DR N-PSDB; AAL62019.
XX PT New human cell adhesion and extracellular matrix proteins (CADECM)
PT polypeptide, useful for preparing a composition for treating a disease
PT associated with decreased expression or overexpression of CADECM e.g.,
PT cancer.
XX Claim 1; Page 279-280; 374pp; English.
XX The invention relates to human cell adhesion and extracellular matrix
XX proteins (CADECM) and nucleic acid molecules encoding such proteins.
XX CADECM proteins are useful for preparing a composition for diagnosing or
XX treating a disease or condition associated with decreased expression or
XX overexpression of functional CADECM e.g., immune disorders or cancer. The
XX invention is also useful in gene therapy. The present sequence is human
XX CADECM protein
XX Sequence 478 AA;
XX Query Match 23.7%; Score 63.5; DB 7; Length 478;
XX Best Local Similarity 40.5%; Pred. No. 22;
XX Matches 17; Conservative 2; Mismatches 20; Indels 3; Gaps 1;
QY 5 PR---PGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
DB 268 PRPGVGTGPMGASGPGYPLRGGRAPSPGYPQSPYPATGKGP 309
RESULT 39
ABO79547
ID ABO79547 standard; protein; 721 AA.
XX AC ABO79547;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #11722.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.

XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI: 2003-615309/58.
XX DR N-PSDB; ABD13118.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 28293; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX Sequence 721 AA;
XX Query Match 23.7%; Score 63.5; DB 7; Length 721;
XX Best Local Similarity 36.8%; Pred. No. 36;
XX Matches 21; Conservative 4; Mismatches 17; Indels 15; Gaps 2;
QY 4 GPRP-----GAPCRAGAPTIVLTSGRRQTLSHG-----SSSPARATLGKPLV 45
DB 242 GRRPDPQAGAPCTARATTGRAGSGRRLTSPGCRWRWTCGRPANACRRSLGSPSV 298
RESULT 40
ABP01580
ID ABP01580 standard; protein; 53 AA.
XX AC ABP01580;
XX DT 24-JUN-2002 (first entry)
XX DE Human ORFX protein sequence SEQ ID NO:3142.
XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
XX degenerative disorder; osteoarthritis; neurodegenerative disorder;
XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
XX hypertension; hypothyroidism; cholesterol ester storage disease;
XX immune deficiency; immune disorder; infectious disease;
XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
XX myasthenia gravis.
XX OS Homo sapiens.
XX PN WO200192523-A2.
XX XX 06-DEC-2001.
XX PD

XX PN EPI347046-A1.
XX PD 24-SEP-2003.
XX PF 12-APR-2002; 2002EP-00008400.
XX PR 22-MAR-2002; 2002JP-00137785.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX DR WPI; 2003-723558/69.
XX DR N-PSDB; ADM01466.
XX PT New polynucleotides and polypeptides are useful in gene therapy, for
XX PT developing a diagnostic marker or medicines for regulating their
XX PT expression and activity, or as a target of gene therapy.
XX PS Claim 1; SEQ ID NO 2594; 305pp; English.
XX CC The invention relates to a novel human polynucleotide and the encoded
XX CC polypeptide. A polynucleotide of the invention may have a use in gene
XX CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
XX CC as a primer for synthesizing the polynucleotide or as a probe for
XX CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX CC useful in gene therapy, for developing a diagnostic marker or medicines
XX CC for regulating their expression and activity, or as a target of gene
XX CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX CC are useful as pharmaceutical agents. The present sequence represents a
XX CC protein sequence of the invention.
XX SQ Sequence 128 AA;
Query Match 23.5%; Score 63; DB 7; Length 128;
Best Local Similarity 35.6%; Pred. No. 5.2;
Matches 16; Conservative 4; Mismatches 23; Indels 2; Gaps 1;
QY 4 GPRGAPCRGAPTIVLTSGRRQTLHSGSSSPARATLGPVLND 48
DB 2 GPAAGHAAPRGPPVPATTAG--AALRAGASEPRQTLGAPLALGE 44
RESULT 43
ABO71209
ID ABO71209 standard; protein; 613 AA.
XX AC ABO71209;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #3384.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX XX

DR WPI; 2003-615309/58.
DR N-PSDB; ABD04780.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 19955; 455pp; English.
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biochip technology. Sequences ABO67826-
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 613 AA;
Query Match 23.5%; Score 63; DB 7; Length 613;
Best Local Similarity 41.0%; Pred. No. 34;
Matches 16; Conservative 5; Mismatches 16; Indels 2; Gaps 1;
QY 5 PRPGAPCRGAPTIVLTSGRRQTLHSGSSSPARATLGP 43
DB 165 PRPGADPRGPARQAARIAGHR--AHGPSANPDAPLGRP 201
RESULT 44
ABO73584
ID ABO73584 standard; protein; 1129 AA.
XX AC ABO73584;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #5759.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD07155.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 22330; 455pp; English.
XX XX

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences AB067826-
CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 1129 AA;

Query Match 23.5%; Score 63; DB 7; Length 1129;
Best Local Similarity 44.8%; Pred. No. 72;
Matches 13; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 13 AGAPTIVLTSGRRQTLHGSSSPARATLG 41
Db 938 SSGCTVLVLSGARIDLRHGAAPAEQVDG 966

RESULT 45

AAU21041

ID AAU21041 standard; protein; 85 AA.

XX AAU21041;

DT 17-DEC-2001 (first entry)

DE Human novel foetal antigen, SEQ ID NO 1285.

KW Human; foetal tissue antigen; antiinflammatory; neuroprotective;
KW immunomodulator; cardiovascular; cystostatic; nephrothropic;
KW cardiovascular; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; breast neoplasm; cancer;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; food additive.

OS Homo sapiens.

FN WO20015312-A2.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001321.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-022824P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 22:54:12 ; Search time 37 Seconds
(without alignments)
135.224 Million cell updates/sec

Title: US-09-819-144A-2

Perfect score: 268

Sequence: 1 MAAGPRGAPCRAGPTIVL.....SSPARATLKGKPLVLDYSLN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: Piri:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	25.4	315	2 C70810	hypothetical prote
2	60.5	22.6	504	2 T33404	hypothetical prote
3	60	22.4	650	2 T48060	beta-D-glucan exoh
4	59	22.0	1887	2 T30176	EGF repeat transme
5	59	22.0	4180	2 G83559	hypothetical prote
6	58.5	21.8	323	2 T19142	hypothetical prote
7	58	21.6	419	2 T36272	hypothetical prote
8	57.5	21.5	886	2 S07132	hypothetical prote
9	57.5	21.5	1611	2 T38236	hypothetical prote
10	57	21.3	308	2 D70875	probable PE protei
11	57	21.3	313	2 A46233	FLAT element-bindi
12	57	21.3	457	2 C70589	probable cobB prot
13	57	21.3	474	2 H86303	hypothetical prote
14	57	21.3	775	2 D86261	hypothetical prote
15	57	21.3	1487	1 ED8EE1	immediate-early pr
16	57	21.3	1487	1 ED8EF6	155K transcription
17	56.5	21.1	136	2 T36547	hypothetical prote
18	56.5	21.1	205	2 C89957	hypothetical prote
19	56.5	21.1	408	2 B87436	conserved hypothet
20	56.5	21.1	640	2 A87715	glucose inhibited
21	56.5	21.1	757	2 T09081	telomere-associate
22	56	20.9	430	2 S24390	transforming prote
23	56	20.9	432	2 T46420	hypothetical prote
24	56	20.9	540	2 S21825	vicilin-like stora
25	56	20.9	616	2 A72627	probable 2-oxoacid
26	56	20.9	780	2 F84470	probable retroelem
27	55.5	20.7	701	2 S61239	hypothetical prote
28	55.5	20.7	1067	2 D75625	probable extracell
29	55.5	20.7	1408	2 S16148	gene serrate prote

30	55.5	20.7	1724	2 T18343	P-glycoprotein - S
31	55	20.5	296	2 A40996	phenylalanine 4-mo
32	55	20.5	323	2 C86384	unknown protein li
33	55	20.5	352	2 JC2466	inhibin beta-C cha
34	55	20.5	364	1 TVHUM1	transforming prote
35	55	20.5	445	2 E87561	hypothetical prote
36	55	20.5	493	2 G84263	long-chain fatty-a
37	55	20.5	513	2 AD1959	ATP-dependent RNA
38	55	20.5	677	2 E70722	hypothetical prote
39	54.5	20.3	251	2 A55523	hypothetical prote
40	54.5	20.3	420	2 AG1859	conserved membrane
41	54.5	20.3	552	2 E87226	trimethylamine N-o
42	54.5	20.3	848	2 D85635	trimethylamine N-o
43	54.5	20.3	848	2 H90772	trimethylamine N-o
44	54.5	20.3	848	2 C64841	trimethylamine N-o
45	54	20.1	178	2 T30717	probable virion pr

ALIGNMENTS

RESULT 1

C70810
hypothetical protein Rv0819 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70810
R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70810
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-315 <COL>
A:Cross-references: UNIPROT:O53831; UNIPARC:UPI00000D119B; GB:AL022004; GB:AL123456; NID:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0819
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0819

Query Match 25.4%; Score 68; DB 2; Length 315;
Best Local Similarity 28.2%; Pred. No. 1.9;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;

QY	1	MAAGPRGAPC-----	-----RAGAPT-----IVLTSGRRQTL 28
DB	51	LVAGSRPGGPIGYLNLSPRGAGMAELVVHPOSRRRGIGTAMARAALAKTAGNQFW 110	
QY	29	SHGSSSPARAT 39	
DB	111	AHGTLDPARAT 121	

RESULT 2

T33404
hypothetical protein H10E21.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33404
R:Davidson, S.; Wohldmann, P.; Courtney, L.
A:Description: The sequence of C. elegans cosmid H10E21.
A:Reference number: Z21339
A:Accession: T33404
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-504 <DAV>
A:Cross-references: UNIPROT:Q8ITW8; UNIPARC:UPI0000084101; EMBL:AF078783; PIDN:AAC26921.1
A:Experimental source: strain Bristol N2; clone H10E21

RESULT 27
S61239
hypothetical protein - bovine herpesvirus 1
C:Species: bovine herpesvirus 1
C:Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004
C:Accession: S61239
R:Vicek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Sch
submitted to the EMBL Data Library, January 1995
A:Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus
A:Reference number: S61233
A:Accession: S61239
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-701 <VLC>
A:Cross-references: UNIPROT:Q65568; UNIPARC:UPI00000F358D; EMBL:Z48053; NID:g9711311; PID
C:Superfamily: varicella-zoster virus gene 43 protein

Query Match 20.7%; Score 55.5; DB 2; Length 701;
Best Local Similarity 36.2%; Pred. No. 1.2e+02;
Matches 21; Conservative 5; Mismatches 19; Indels 13; Gaps 3;

QY 2 AAGPRPGACRA-GAPTIVLTSGRRQ--TLGH-----GSSSPARATLKGPLVL 46
DB 325 AAGPLGAVARALAGDTLVATAATQATLTWFPDAGHCALLGGTTPCDWRRRLSL 382

RESULT 28
D75625
probable extracellular nuclease - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: D75625
R:White, O.; Eiken, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75625
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1067 <WHI>
A:Cross-references: UNIPROT:Q9RZQ3; UNIPARC:UPI00000D3B4A; GB:AE001826; NID:g6460827; PI
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0067
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MPI

Query Match 20.7%; Score 55.5; DB 2; Length 1067;
Best Local Similarity 31.2%; Pred. No. 1.9e+02;
Matches 15; Conservative 9; Mismatches 23; Indels 1; Gaps 1;

QY 2 AAGPRPGACRA-GAPTIVLTSGRRQTLGHGSSSPARATLKGPLVNDY 49
DB 875 AAGTVPVTTTGGAGLTATATLNTVSVGGTPPSST-GSDLIIIEY 921

RESULT 29
S16148
gene serrate protein precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1991 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C:Accession: S16148; S16878; A36666
R:Thomas, U.; Speicher, S.A.; Knust, E.
Development 111, 749-761, 1991
A:Title: The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a co
A:Reference number: S16148; MUID:91347903; PMID:1840519
A:Accession: S16148
A:Molecule type: mRNA
A:Residues: 1-1408 <THO1>

A:Cross-references: UNIPROT:P18168; UNIPARC:UPI000017A1AA; EMBL:X56811
R:Thomas, U.
submitted to the EMBL Data Library, November 1990
A:Reference number: S16878
A:Accession: S16878
A:Molecule type: mRNA
A:Residues: 1-1351, 'T', 1353-1408 <THO2>
A:Cross-references: UNIPARC:UPI000013588E; EMBL:X56811; NID:g8563; PID:g8564
R:Fleming, R.J.; Scottgale, T.N.; Diederich, R.U.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2188-2201, 1990
A:Title: The gene Serrate encodes a putative EGF-like transmembrane protein essential fo
A:Reference number: A36666; MUID:91099666; PMID:2125287
A:Accession: A36666
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-15, 20-26, 'A', 28-1408 <FILE>
A:Cross-references: UNIPARC:UPI00000498BA; GB:M35759; NID:g158605; PID:g158606
C:Genetics:
A:Gene: FlyBase:Ser
A:Cross-references: FlyBase:FBgn0004197
C:Keywords: glycoprotein; transmembrane protein
F:1-84/Domain: signal sequence #status predicted <SIG>
F:85-1408/Product: gene serrate protein #status predicted <MAT>
F:85-1221/Domain: extracellular #status predicted <EXT>
F:283-316/Domain: EGF homology <EG01>
F:319-348/Domain: EGF homology <EG02>
F:355-388/Domain: EGF homology <EG03>
F:395-488/Domain: EGF homology #status atypical <EG04>
F:495-526/Domain: EGF homology <EG05>
F:533-608/Domain: EGF homology #status atypical <EG06>
F:615-645/Domain: EGF homology <EG07>
F:652-683/Domain: EGF homology <EG08>
F:690-720/Domain: EGF homology <EG09>
F:727-796/Domain: EGF homology #status atypical <EG10>
F:803-834/Domain: EGF homology <EG11>
F:841-876/Domain: EGF homology <EG12>
F:883-914/Domain: EGF homology <EG13>
F:921-952/Domain: EGF homology <EG14>
F:997-1060/Region: cysteine-rich
F:1222-1246/Domain: transmembrane #status predicted <TM1>
F:1247-1408/Domain: intracellular #status predicted <INT>
F:152,196,247,331,412,452,558,739,965,977,1004,1030,1150/Binding site: carbohydrate (Asn)

Query Match 20.7%; Score 55.5; DB 2; Length 1408;
Best Local Similarity 35.9%; Pred. No. 2.5e+02;
Matches 14; Conservative 5; Mismatches 11; Indels 9; Gaps 2;

QY 10 PCRAGPTIVLTSGRRQTLGHGSSSPARATLKGPL 44
DB 399 PCRNGTCTLTSTNRTQAQVTRTSHG-----RSNMGRPV 432

RESULT 30
T18343
P-glycoprotein - Sauroleishmania tarentolae
C:Species: Sauroleishmania tarentolae
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18343
R:Legare, D.; Hettema, E.; Ouellette, M.
Mol. Biochem. Parasitol. 68, 81-91, 1994
A:Title: The P-glycoprotein-related gene family in Leishmania.
A:Reference number: Z18879; MUID:95198776; PMID:7891750
A:Accession: T18343
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1724 <LEG>
A:Cross-references: UNIPROT:Q25425; UNIPARC:UPI000007AAA07; EMBL:L29485; NID:g460312; PID:
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 20.7%; Score 55.5; DB 2; Length 1724;
Best Local Similarity 35.0%; Pred. No. 3e+02;
Matches 21; Conservative 5; Mismatches 21; Indels 13; Gaps 4;

QY 5 PRPG-APCRAGAPTIVLTSGRRQTLSHGS-----SSPARATLGKPLVIND---YSIN 52
Db 713 PSPSTAPSASADPAWEPLTTERNTESHGDAASPEPLSSPAH-TYNSEVVFNEDSTYILN 771

RESULT 31
A40996
phenylalanine 4-monoxygenase (EC 1.14.16.1) - Chromobacterium violaceum
N/Alternate names: phenylalaninase; phenylalanine 4-hydroxylase
C/Species: Chromobacterium violaceum
C/Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004
A/Accession: A40996
R/Onishi, A.; Liotta, L.J.; Benkovic, S.J.
J. Biol. Chem. 266, 18454-18459, 1991
A/Title: Cloning and expression of Chromobacterium violaceum phenylalanine hydroxylase
A/Reference number: A40996; MUID:92011593; PMID:1655752
A/Molecule type: DNA
A/Accession: A40996
A/Residues: 1-296 <ONI>
A/Cross-references: UNIPROT:P30967; UNIPARC:UPI000016BBD9; GB:M55915; NID:g144481; PIDN:
A/Accession: B40996
A/Molecule type: protein
A/Residues: 1-20 <ONS>
A/Cross-references: UNIPARC:UPI000017AA1E
A/Experimental source: strain ATCC 12540
C/Keywords: bioplerin; oxidoreductase

Query Match 20.5%; Score 55; DB 2; Length 296;
Best Local Similarity 38.9%; Pred. No. 59;
Matches 14; Conservative 5; Mismatches 11; Indels 6; Gaps 1;

QY 12 RAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLN 47
Db 254 RAGFRSAILAAGRATVGRGDIAP-----DDLVLN 283

RESULT 32
C86384
unknown protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
A/Accession: C86384
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Huggins, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: C86384
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-323 <STO>
A/Cross-references: UNIPROT:Q9C6L2; UNIPARC:UPI00000AC544; GB:AE005172; NID:gi00923333; F
C/Genetics:
A/Map position: 1

Query Match 20.5%; Score 55; DB 2; Length 323;
Best Local Similarity 32.6%; Pred. No. 64;
Matches 15; Conservative 8; Mismatches 15; Indels 8; Gaps 1;

QY 1 MAAGPRGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVL 46
Db 285 LASADATATAARAAQVLV-----NVSYGATTPTVAABGKPLNL 322

RESULT 33
JC2466
inhibin beta-C chain precursor - human

N/Alternate names: activin beta C chain
C/Species: Homo sapiens (man)
C/Date: 29-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C/Accession: JC2466
R/Hoetten, G.; Neidhardt, H.; Schneider, C.; Pohl, J.
Biochem. Biophys. Res. Commun. 206, 608-613, 1995
A/Title: Cloning of a new member of the TGF-beta family: A putative new activin betaC chi
A/Reference number: JC2466; MUID:95126961; PMID:7826378
A/Molecule type: mRNA
A/Residues: 1-352 <HOE>
A/Cross-references: UNIPROT:P55103; UNIPARC:UPI000012D42A; GB:X82540; NID:g669154; PIDN:
A/Experimental source: liver
C/Genetics:
A/Gene: GDB:INHBC
A/Cross-references: GDB:632884
A/Map position: 2cen-2q13
C/Superfamily: inhibin
C/Keywords: glycoprotein
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-236/Domain: propeptide #status predicted <PRO>
F/237-352/Product: activin beta C #status predicted <MAT>
F/110,143,161/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.5%; Score 55; DB 2; Length 352;
Best Local Similarity 36.0%; Pred. No. 70;
Matches 18; Conservative 3; Mismatches 21; Indels 8; Gaps 2;

QY 3 AGPRGAPCRGA-CAPTIVLTSGRRQTLSHGS-----PARATLGKPL 44
Db 18 ATPRAGQCPCAGPTLELESQRELLDLAKRSILDKHLTQRTPLNRPV 67

RESULT 34
TVHUM,
transforming protein L-myc-1 - human
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004
C/Accession: A27675; S01200; S09390
R/Kaye, F.; Battey, J.; Nau, M.; Brooks, B.; Seifter, E.; De Greve, J.; Birrer, M.; Saus,
Mol. Cell. Biol. 8, 186-195, 1988
A/Title: Structure and expression of the human L-myc gene reveal a complex pattern of alt
A/Reference number: A27675; MUID:88094386; PMID:2827002
A/Accession: A27675
A/Molecule type: DNA
A/Residues: 1-364 <KAY>
A/Cross-references: UNIPROT:P12524; UNIPARC:UPI000012FAF3; GB:M19720; NID:g188906; PIDN:
R/Depinho, R.A.; Hattori, K.S.; Tefsaye, A.; Yancopoulos, G.D.; Alt, F.W.
Genes Dev. 1, 1311-1326, 1987
A/Title: The human myc gene family: structure and activity of L-myc and an L-myc pseudoge
A/Reference number: S01200; MUID:88112807; PMID:3322939
A/Accession: S01200
A/Molecule type: DNA
A/Residues: 1-364 <DEP>
A/Cross-references: UNIPARC:UPI000012FAF3
R/Ikegaki, N.; Minna, J.; Kennett, R.H.
EMBO J. 8, 1793-1799, 1989
A/Title: The human L-myc gene is expressed as two forms of protein in small cell lung ca
A/Reference number: S09390; MUID:89356654; PMID:2548855
A/Accession: S09390
A/Status: preliminary
A/Molecule type: protein
A/Residues: 165-179 <IKE>
A/Cross-references: UNIPARC:UPI00001732D0
C/Genetics:
A/Gene: GDB:MYCL1
A/Cross-references: GDB:1120706; OMIM:164850
A/Map position: 1p32-1p32
A/Introns: 166/1
C/Superfamily: myc transforming protein; myc transforming protein homology
C/Keywords: DNA binding; nucleus; transforming protein
F/1-364/Domain: myc transforming protein homology <MYC>

```
Query Match      20.5%; Score 55; DB 1; Length 364;
Best Local Similarity 31.1%; Pred. No. 72;
Matches 19; Conservative 5; Mismatches 13; Indels 24; Gaps 4;

QY 2 AAGPRGAPCRAGAPTIVLTSGRRQTLSHGSSSPA-----RATLG--KPLV 45
    |||||
DB 137 AGNPAPAPCPPLGEP-----KTAQCS-GSESFSDSENEIDVVYVKRQSLGIRKPT 188
    |||||
QY 46 L 46
    :
DB 189 I 189

RESULT 35
E87561
hypothetical protein CC2518 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: E87561
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: E87561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-445 <STO>
A:CROSS-references: UNIPROT:Q9ASD2; UNIPARC:UPI00000C7762; GB:AE005673; NID:g13424075; E
C:Genetics:
A:Gene: CC2518

Query Match      20.5%; Score 55; DB 2; Length 445;
Best Local Similarity 43.2%; Pred. No. 88;
Matches 19; Conservative 3; Mismatches 18; Indels 4; Gaps 2;

QY 1 MAAGPRGAPCRAG---APTIVLTSGRRQTLSHGSSSPARATLG 41
    :|||
DB 252 LAGGPFDEAAATGVQARMITLLAQGRQTFTNGWSSP-RANLG 294
    :|||

RESULT 36
G84263
long-chain fatty-acid-CoA ligase [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: G84263
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithäuser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: AB4160; MUID:20504483; PMID:11016950
A:Accession: G84263
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-493 <STO>
A:CROSS-references: UNIPROT:Q9HQH8; UNIPARC:UPI00000637E1; GB:AE004437; NID:g10580623; E
C:Genetics:
A:Gene: lfl1
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match      20.5%; Score 55; DB 2; Length 493;
Best Local Similarity 34.0%; Pred. No. 98;
Matches 18; Conservative 7; Mismatches 12; Indels 16; Gaps 4;

QY 8 GAP-----CRAGAPTIVLTSGRRQTLSS-----HGSSSPAR-ATLGKPLV 45
    |||||
DB 268 GAPTPPPELVACAGVP-VCPTYGATETASQATLHASDAPAHPSVGKPL 319
    |||||
```

```
RESULT 37
AD1959
ATP-dependent RNA helicase [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AD1959
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasa moto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD1959
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-513 <KUR>
A:CROSS-references: UNIPROT:Q8YXJ0; UNIPARC:UPI00000CE00E; GB:BA000019; PIDN:BA073180.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1223

Query Match      20.5%; Score 55; DB 2; Length 513;
Best Local Similarity 40.6%; Pred. No. 1e+02;
Matches 13; Conservative 5; Mismatches 10; Indels 4; Gaps 1;

QY 12 RAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
    |||||
DB 479 RRGTPKPKLTGRRET----SATPSNPKLGSP 506
    |||||

RESULT 38
E70722
hypothetical protein Rv1836c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70722
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome :
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70722
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-677 <COL>
A:CROSS-references: UNIPROT:Q50597; UNIPARC:UPI000013B324; GB:Z78020; GB:AL123456; NID:g1
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1836c
C:Superfamily: Mycobacterium leprae hypothetical protein MLCB1788.28

Query Match      20.5%; Score 55; DB 2; Length 677;
Best Local Similarity 38.1%; Pred. No. 1.4e+02;
Matches 16; Conservative 3; Mismatches 23; Indels 0; Gaps 0;

QY 2 AAGPRGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
    |||||
DB 317 AASAPAGAPATAGICAVRTLMGARKPLADDSLTAAAMDTLLKP 358
    |||||

RESULT 39
A55523
hypothetical protein (bphD 3' region) - Pseudomonas sp. (strain KKS102)
C:Species: Pseudomonas sp.
C>Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 11-Jan-2002
C:Accession: A55523
R:Kikuchi, Y.; Nagata, Y.; Hinata, M.; Kimbara, K.; Fukuda, M.; Yano, K.; Takagi, M.
J. Bacteriol. 176, 1689-1694, 1994
A:Title: Identification of the bphA4 gene encoding ferredoxin reductase involved in bioph
A:Reference number: A55523; MUID:94179104; PMID:8132464
A:Accession: A55523
```

RESULT 42

D85635
trimethylamine N-oxide reductase subunit [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85635
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lin, A.; Dimailanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: D85635
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-848 <STO>
A:Cross-references: UNIPARC:UPI0000165754; GB:AE005174; NID:g12514264; PIDN:AAG55544.1; C
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: torA
C:Superfamily: trimethylamine-N-oxide reductase

Query Match 20.3%; Score 54.5; DB 2; Length 848;
Best Local Similarity 36.1%; Pred. No. 1.9e+02;
Matches 13; Conservative 6; Mismatches 10; Indels 7; Gaps 1;

Qy 4 GPRFGAPCRGAPTIVLTSGRRQTLSHGSSSPARAT 39
Db 778 GGEPGALCKYGNPNVL-----TIDIGTSQLAQAT 806

RESULT 43

H90772
trimethylamine N-oxide reductase subunit [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90772
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90772
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-848 <HAY>
A:Cross-references: UNIPROT:P58360; UNIPARC:UPI0000137172; GB:BA000007; PIDN:BAB34575.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECel152
C:Superfamily: trimethylamine-N-oxide reductase

Query Match 20.3%; Score 54.5; DB 2; Length 848;
Best Local Similarity 36.1%; Pred. No. 1.9e+02;
Matches 13; Conservative 6; Mismatches 10; Indels 7; Gaps 1;

Qy 4 GPRFGAPCRGAPTIVLTSGRRQTLSHGSSSPARAT 39
Db 778 GGEPGALCKYGNPNVL-----TIDIGTSQLAQAT 806

RESULT 44

C64841
trimethylamine-N-oxide reductase (EC 1.6.6.9) precursor - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: C64841; S43698; S34222
R:Blattner, P.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: AC4720; MUID:97426617; PMID:9278503

A:Accession: C64841
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-848 <BLAT>
 A:Cross-references: UNIPROT:P33225; UNIPARC:UPI0000137173; GB:AE000201; GB:U00096; NID:9
 A:Experimental source: strain K-12, substrain MG1655
 R:Mejean, V.; Tobbi-Nivol, C.; Lepelletier, M.; Giordano, G.; Chippaux, M.; Pascal, M.C.
 Mol. Microbiol. 11, 1169-1179, 1994
 A:Title: TMAO anaerobic respiration in Escherichia coli: involvement of the tor operon.
 A:Reference number: S43697; MUID:94293785; PMID:8022286
 A:Accession: S43698
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-172, 'R', 174-175, 'R', 177-255, 'R', 257, 'S', 259-280, 'G', 282-324, 'E', 326-347, 'S'
 A:Cross-references: UNIPARC:UPI000016F536; EMBL:X73888; NID:9556701; PIDN:CAAS2095.1; PI
 A:Experimental source: strain K-12
 C:Genetics:
 A:Gene: torA
 C:Function:
 A:Description: reduces trimethylamine-N-oxide into trimethylamine
 C:Superfamily: trimethylamine-N-oxide reductase
 C:Keywords: molybdenum; NAD; oxidoreductase; periplasmic space
 P:1-39/Domain: signal sequence #status predicted <SIG>
 F:40-848/Product: trimethylamine-N-oxide reductase #status predicted <MAT>

Query Match 20.3%; Score 54.5; DB 2; Length 848;
 Best Local Similarity 36.1%; Pred. No. 1.9e+02;
 Matches 13; Conservative 6; Mismatches 10; Indels 7; Gaps 1;

QY 4 GPRGAPCRAGACTIVLTSGRQTLSHGSSSPARAT 39
 DB 778 GEEPGALCKYGNPNVL-----TIDIGTSLQAQAT 806

RESULT 45
 T30717
 Probable virion protein 115L - Molluscum contagiosum virus 1
 N:Alternate names: MC115L
 C:Species: Molluscum contagiosum virus 1
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C:Accession: T30717
 R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
 Science 273, 813-816, 1996
 A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
 A:Reference number: Z20876; MUID:96325459; PMID:8670425
 A:Accession: T30717
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-178 <SEN>
 A:Cross-references: UNIPROT:Q98282; UNIPARC:UPI00000P6094; EMBL:U60315; NID:g1491943; PI
 C:Genetics:
 A:Note: MC115L
 C:Superfamily: vaccinia virus 20K virion protein

Query Match 20.1%; Score 54; DB 2; Length 178;
 Best Local Similarity 29.4%; Pred. No. 46;
 Matches 15; Conservative 10; Mismatches 22; Indels 4; Gaps 2;

QY 2 AAGPRGAPCRAGACTIVLTSGRQTLSHGSSSPARATGKPLVLDYSLN 52
 DB 132 SSSPRPRAP-RSGTRR---STGTRSSGARGSRGSRPRATGDSFSPMDYGM 178

Search completed: December 27, 2005, 23:05:00
 Job time : 39 secs

...the Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 22:53:57 ; Search time 227 Seconds
(without alignments)
161.619 Million cell updates/sec

Title: US-09-819-144A-2

Perfect score: 268

Sequence: 1 MAAGPRGAPCRAGATTIVL.....SSPARATLKGPLVLDYSLN 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	73.5	27.4	290	Q68D58 HUMAN	Q68D58 homo sapien
2	73.5	27.4	350	Q641Q8 HUMAN	Q641Q8 homo sapien
3	73.5	27.4	863	1 MILK1 HUMAN	Q8n3f8 homo sapien
4	72	26.9	636	Q6AYD5 RAT	Q6ayd5 rattus norv
5	71.5	26.7	1581	Q4SA50 TETNG	Q4sa50 tetraodon n
6	70	26.1	1004	Q4P5V3 USTMA	Q4p5v3 ustilago ma
7	68	25.4	315	Q7U173 MYCBO	Q7u173 mycobacteri
8	68	25.4	315	Q53831 MYCTU	Q53831 mycobacteri
9	67.5	25.2	391	Q98A00 RHIZLO	Q98a00 rhizobium l
10	67	25.0	563	Q4UVP4 XANCP	Q4uvp4 xanthomonas
11	67	25.0	563	Q8P8E3 XANCP	Q8p8e3 xanthomonas
12	66.5	24.8	1787	Q9M4X9 CHLRE	Q9m4x9 chlamydomon
13	66	24.6	272	Q743X3 MYCPA	Q743x3 mycobacteri
14	64.5	24.1	215	Q81V50 HUMAN	Q81v50 homo sapien
15	64.5	24.1	596	Q82NB1 STRAW	Q82nb1 streptomyce
16	64	23.9	713	Q4WQ64 ASPFU	Q4wq64 aspergillus
17	64	23.9	631	Q4Q5C4 LEIMA	Q4q5c4 leishmania
18	63.5	23.7	355	Q539T2 PONPY	Q539t2 pongo pygma
19	63.5	23.7	377	Q8N3K4 HUMAN	Q8n3k4 homo sapien
20	63.5	23.7	562	Q8BQZ7 MOUSE	Q8bqz7 mus musculu
21	63.5	23.7	874	Q6PDL7 MOUSE	Q6pdl7 mus musculu
22	63	23.5	128	Q8N0Y6 HUMAN	Q8n0y6 homo sapien
23	63	23.5	171	Q6K3G1 ORYSA	Q6k3g1 oryza sativ
24	63	23.5	206	Q14897 HUMAN	Q14897 homo sapien
25	63	23.5	236	Q5QPLO HUMAN	Q5qplo homo sapien
26	63	23.5	360	1 YBOX2 MOUSE	Q9zzc8 mus musculu
27	63	23.5	373	Q7WNE3 BORBR	Q7wne3 bordetella
28	63	23.5	2063	Q4Q578 LEIMA	Q4q578 leishmania
29	62.5	23.3	150	Q84ZQ6 ORYSA	Q84zq6 oryza sativ
30	62.5	23.3	208	Q5Z9G0 ORYSA	Q5z9g0 oryza sativ
31	62.5	23.3	556	Q8XQ41 RALUSO	Q8xq41 raistonia s

32	62.5	23.3	671	2 Q9Z5X4 FRASE	Q9z5x4 frankia sp.
33	62.5	23.3	771	2 Q5ZKM6 CHICK	Q5zkm6 gallus gall
34	62	23.1	215	2 Q93392 COTCO	Q93392 cturnix co
35	62	23.1	259	2 Q6ZM28 HUMAN	Q6zm28 homo sapien
36	62	23.1	415	2 Q6A047 MOUSE	Q6a047 mus musculu
37	62	23.1	449	2 Q7SGS8 NEUCR	Q7sgs8 neurospora
38	62	23.1	496	2 Q7U8L8 SYNXP	Q7u8l8 synechococc
39	62	23.1	1050	2 Q4T9N9 TETNG	Q4t9n9 tetraodon n
40	62	23.1	1059	2 Q5KG17 CRYNE	Q5kg17 cryptococcu
41	62	23.1	1059	2 Q5S585 CRYNE	Q5s585 cryptococcu
42	62	23.1	1325	1 RIMB2 CHICK	Q8qfx1 gallus gall
43	62	23.1	1553	2 Q4SE53 TETNG	Q4se53 tetraodon n
44	61.5	22.9	588	2 Q9N2G7 RABIT	Q9n2g7 oryctolagus
45	61.5	22.9	1178	2 Q69ZH8 MOUSE	Q69zh8 mus musculu

ALIGNMENTS

RESULT 1					
Q68D58 HUMAN					
ID	Q68D58	HUMAN	PRELIMINARY;	PRT;	290 AA.
AC	Q68D58;				
DT	25-OCT-2004	(Tremblrel. 28, Created)			
DT	25-OCT-2004	(Tremblrel. 28, Last sequence update)			
DT	25-OCT-2004	(Tremblrel. 28, Last annotation update)			
DE	Hypothetical protein DKFZp686M2226.				
GN	Name=DKFZp686M2226;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;				
OC	Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISSUE=Testis;				
RG	The German cDNA Consortium;				
RA	Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,				
RA	Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;				
RL	Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL;	CR749566;	CAH18362.1;	-;	mRNA.
KW	Hypothetical protein.				
SQ	SEQUENCE 290 AA; 32971 MW; DF4D273FB8C84BCB CRC64;				
Query Match 27.4%; Score 73.5; DB 2; Length 290;					
Best Local Similarity 45.7%; Pred. No. 2.5;					
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;					
QY	2	AAGPRFGAPCRAGATTIVLTSRRQTL-SHGSSSP	35		
Db	15	SSGPQAKPCSGATPTLLLVGDRSPVPSPGSSSP	49		

```
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL proxy Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2].
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RG NIH MGC Project;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082243; AAB2243.1; -, mRNA.
FT NON-ITER
SQ
Query Match 27.4%; Score 73.5; DB 2; Length 350;
Best Local Similarity 45.7%; Pred. No. 3.1;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

Qy 2 AAGPRGAPCAGACTIVLTSGRRQTL-SHGSSSP 35
Db 75 SSGPQAPKPCSGATPTLLVGDSPVSPGSSSP 109

RESULT 3
MILK1_HUMAN STANDARD; PRT; 863 AA.
AC Q8N3F8; Q5T116; Q7RTP5; Q8N3N8; Q9BVL9; Q9BY92; Q9UH43; Q9UH44;
AC Q9UH45;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Molecule interacting with Rabi3 (MIRab13) (MICAL-like protein 1).
GN Name=MIRAB13; Synonyms=KIAA1668;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1].
RP NUCLEOTIDE SEQUENCE
RC TISSUE=Uterine adenocarcinoma;
RA Pandjaitan R., Zahroul A.;
RT "MIRab13, a novel effector of Rabi3.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2].
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]
RX PubMed=15461802; DOI=10.1186/gb-2004-5-10-r84;
RA Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,
RA Cole C.G., Goward M.E., Aguado B., Malliya M., Mokrab Y., Huckle E.J.,
RA Beare D.M., Dunham I.;
RT "A genome annotation-driven approach to cloning the human ORFeome.";
RL Genome Biol. 5:RESEARCH84.1-RESEARCH84.11 (2004).
RN [3].
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
```


RT and mouse cDNA sequences.":
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [7]
RP GENE STRUCTURE.
RX MEDLINE=22105479; PubMed=12110185; DOI=10.1016/S0092-8674(02)00794-8;
RA Terman J.R., Mao T., Pasterkamp R.J., Yu H.-H., Kolodkin A.L.;
RT "MICALs, a family of conserved flavoprotein oxidoreductases, function
in plexin-mediated axonal repulsion.";
RL Cell 109:887-900 (2002).
RN [8]
RP PHOSPHORYLATION SITE SER-578.
RX PubMed=15302935; DOI=10.1073/pnas.0404720101;
RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
RT "Large-scale characterization of HeLa cell nuclear phosphoproteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135 (2004).
CC -!- FUNCTION: May be a cytoskeletal regulator.
CC -!- SUBUNIT: Binds to Rab13.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 CH (calponin-homology) domain.
CC -!- SIMILARITY: Contains 1 LIM zinc-binding domain.
CC -!- CAUTION: Ref.4 (CAD39036) sequence differs from that shown due to
frameshifts in positions 486 and 507.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR EMBL; AJ496196; CAD42713.1; -; mRNA.
DR EMBL; CA456437; CAG30323.1; -; mRNA.
DR EMBL; AL022311; CA118864.1; -; Genomic DNA.
DR EMBL; AL834373; CAD39036.1; ALT_FRAME; mRNA.
DR EMBL; AL833860; CAD38718.1; -; mRNA.
DR EMBL; AB051455; BAB33338.1; -; mRNA.
DR EMBL; BC001090; AAH01090.2; -; mRNA.
DR EMBL; BK000466; DAA01345.1; -; mRNA.
DR HSSP; Q01082; 1BRK.
DR Ensembl; ENSG00000100139; Homo sapiens.
DR GO; GO:0005622; C:intracellular; IC.
DR GO; GO:0016192; P:vesicle-mediated transport; IC.
DR InterPro; IPR001715; Calponin_act_bd.
DR InterPro; IPR001781; LIM_Zn_bd.
DR Pfam; PF00307; CH; 1.
DR Pfam; PF00412; LIM; 1.
DR ProDom; PD000094; LIM; 1.
DR SMART; SM00033; CH; 1.
DR SMART; SM00132; LIM; 1.
DR PROSITE; PS50021; CH; 1.
DR PROSITE; PS00478; LIM DOMAIN 1; 1.
DR PROSITE; PS00023; LIM DOMAIN 2; 1.
KW Coiled coil; Cytoskeleton; LIM domain; Metal-binding; Phosphorylation;
KW Polymorphism; Zinc.
FT DOMAIN 2 105 CH.
FT DOMAIN 162 225 LIM zinc-binding.
FT COILED 682 711 Potential.
FT COILED 785 830 Potential.
FT COMPIAS 253 667 Pro-rich.
FT COMPIAS 428 437 Poly-Glu.
FT MOD_RES 578 Phosphoserine.
FT VARIANT 519 A -> S (in dbSNP:9610875).
FT VARIANT 583 /FTID=VAR_018262.
FT VARIANT 593 P -> L (in dbSNP:2272829).
FT CONFLICT 137 P -> S (in Ref. 5).
FT CONFLICT 212 F -> Y (in Ref. 4; CAD39036).
SQ SEQUENCE 863 AA; 93441 MW; 755E3B57C6037292 CRC64;

Query Match 27.4%; Score 73.5; DB 1; Length 863;
Best Local Similarity 45.7%; Pred. No. 8.7;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 2 AAGPRGAPCAGACTVLTSGRQTL-SHGSSSP 35
DB 588 SSGPQAPKPCGATPTLLLVGDRSPVPSPGSSSP 622
RESULT 4
Q6AYD5 RAT PRELIMINARY; PRT; 636 AA.
ID Q6AYD5; AC Q6AYD5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE G1 to S phase transition 1.
GN Name=Septi;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC079092; AAH79092.1; -; mRNA.
DR GO; GO:0005622; C:intracellular; ISS.
DR GO; GO:0003924; P:GTPase activity; ISS.
DR GO; GO:0005515; P:protein binding; ISS.
DR GO; GO:0003747; P:translation release factor activity; ISS.
DR GO; GO:0000082; P:G1/S transition of mitotic cell cycle; ISS.
DR GO; GO:0000184; P:mRNA catabolism, nonsense-mediated decay; ISS.
DR InterPro; IPR009818; Ataxin-2 C.
DR InterPro; IPR004160; EFTU Cterm.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; ProtSyn_GTPbind.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR Pfam; PF07145; FAM2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACITOR GTP; UNKNOWN 1.
SQ SEQUENCE 636 AA; 68751 MW; 6FB5F124F5132D2C CRC64;

Query Match 26.9%; Score 72; DB 2; Length 636;
Best Local Similarity 39.1%; Pred. No. 9.3;
Matches 18; Conservative 7; Mismatches 21; Indels 0; Gaps 0;

QY 2 AAGPRGAPCAGACTVLTSGRQTL-SHGSSSPARATLGKPLVN 47
DB 588 SSGPQAPKPCGATPTLLLVGDRSPVPSPGSSSP 622

Db 101 AAGD HGAGSSAGGPPEPVESQEQSLCEGSNSTVMELSEPVEN 146

RESULT 5

Q4SA50_TETNG
ID Q4SA50_TETNG PRELIMINARY; PRT; 1581 AA.
AC Q4SA50_TETNG
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 12 SCAP14692, whole genome shotgun sequence.

OS	taxname=Salmonella021897001;
OS	Tetradon nigriviridis (Green puffer).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC	Tetraodontidae; Tetraodontidae; Tetradon.
OC	NCBI_TaxID=99883;
RN	[1]

NU
[T]
NP
NUCLEOTIDE SEQUENCE

JAILLON O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
R. Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
R. Dasilva C., Salanoubat M., Levy M., Boudet N., Castellan S.,
R. Anthonard V., Jubin C., Castellì V., Katinka M., Vacherie B.,
R. Blemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
R. Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
R. Parra G., Lardier G., Chaplie C., McKernan K.J., McEwan P., Bosak S.,
R. Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
R. Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
R. Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
R. Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
R. "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
R. the early vertebrate proto-karyotype.";
R. Nature 431:946-957(2004).

RN [2]

RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; CAAE0104692; CAG02482.1; -; Genomic_DNA.
DR InterPro; IPR07108; Cut homeo.
DR InterPro; IPR003350; Hmoec CUT.
DR InterPro; IPR001356; Homeobox.
DR

DR pfam: PF02376; CUT: 3.

DR PFam: PF000046; Homeobox, 1.
DR PDom: PD000010; Homeobox, 1.
DR SMART: SM00389; HOX, 1.
DR PROSITE: PS10442; CUT, 3.
DR PROSITE: PS50071; HOMEBOX 2, 1.
DR DNA-binding: Homeobox; Nuclear protein.
DR SK SEQUENCE 1591 AA; 170831 MW; 710959B08CC7597A CRC64;
DR

Query Match	26.7%	Score	71.5	DB	2	Length	1581
Best Local Similarity	44.2%	Pred.	No. 30				
Matches	19	Conservative	9	Mismatches	12	Indels	3
						Gaps	2

Qy		1	MAAGPRPGA--PCRAGAPTIVLTSGRRQTLSHGS-SSPARATL 40 : : : : : : : : : : : : :
Dd		1065	MLSRPFPGASSPRKAGSPSSVCGYWTSSARASASTPARATL 110 : : : : : : : : : : : : : :

RESULT 6

Q4P5V3 USTWA
ID Q4P5V3 USTWA PRELIMINARY; PRT; 1004 AA.
AC Q4P5V3;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Hypothetical protein.

GN ORFNames=UM04510.1:

OS · *Ustilago maydis* 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI TaxID=237631;

RN [1]

RC NUCLEOTIDE SEQUENCE.

RP STRAIN=521.

RA Birren B., Numbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,
RA Rayul T., Blishteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgaltay B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley C., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamyselis M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira K., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabellia R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrum J., Menues L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotono B.,
RA O'Neill K., Oman S., Parker S., Perrin D., Phunthang P., Pignani B.,
RA Purcell S., Rachupka T., Ramaeamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tensing P., Tsefaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliou H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;

RT "The genome sequence of *Ustilago maydis*.";

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

RC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data

CC EMBL; ACP01000160; EAK85392.1; -; Genomic_DNA.

DR Hypothetical protein.

SW SEQUENCE 1004 AA; 108726 MW; 3948B22248CF953A CRC64

Query Match 26 1%: score 70: DB 2: Length 1004:

Gap, match	Score 10; 25	Length 1001;
Best Local Similarity	42.28;	
Pred. No. 27;		
Best Local Similarity	42.6%;	
Matches 20; Conservative	2; Mismatches	21; Indels 4; Gaps 1;

Qy . . . 3 AGPRPGACRAGAPTIVLT-----SGRRQTLSHGSSSPARATLGKPLV 45
| | | | | : | | | | |
Dd . . . 539 AAPRNVDCEALPQTAPLTLTKQATRSSLSDKDDSSPPRTTRAPPV 585

RESULT 7

Q7U173	MYCBO	PRELIMINARY;	PRT;	315 AA.
AC	Q7U173	MYCBO	PRELIMINARY;	
ID	Q7U173			
AD	Q7U173;			
DT	01-OCT-2003	(TrEMBLrel. 25, Created)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE	Hypothetical protein Mb0842.			
GN	OrderedLocusNames=Mb0842;			

[illegible]

Db 211 KAGADMRSRG---IITSGSRNTLHGGSSNNRRLRGDPV 246

RESULT 10

Q4UVPA_XANCP PRELIMINARY; PRT; 563 AA.
 ID Q4UVPA_XANCP PRELIMINARY; PRT; 563 AA.
 AC Q4UVPA;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE N-acetylmuramoyl-L-alanine amidase.
 GN ORFNames=XC_1816;
 OS Xanthomonas campestris pv. campestris str. 8004.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=314565;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=8004;
 RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
 RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
 RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
 RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
 RA He C.-Z.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CP000050; AAY48879.1; -; Genomic DNA.
 SQ SEQUENCE 563 AA; 57271 MW; 5860B59189C35941 CRC64;

Query Match 25.0%; Score 67; DB 2; Length 563;
 Best Local Similarity 46.9%; Pred. No. 31;
 Matches 15; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

Qy 5 PRPGAPCRAGACTIVLTSGRRQTLSHGSSSPA 36

Db 229 PRPGTVAAPGAP--ATTGGTRTTIATGTPSPA 258

RESULT 11

Q8P8E3_XANCP PRELIMINARY; PRT; 563 AA.
 ID Q8P8E3_XANCP PRELIMINARY; PRT; 563 AA.
 AC Q8P8E3;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE N-acetylmuramoyl-L-alanine amidase.
 GN Name=amiC;
 OS Xanthomonas campestris (pv. campestris)
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RA MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 RA da Silva A.C.R., Ferro J.A., Reibach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
 RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
 RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
 RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L., Cursino-Santos J.R.,
 RA El-Dorzy H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
 RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locati E.C., Machado M.A., Madeiro A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira L.A.F., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinoia L.A.F., Taktai M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.P.,
 RA Setubal J.C., Kitajima J.P.;
 RL "Comparison of the genomes of two Xanthomonas pathogens with differing host specificities";
 RT Nature 417:459-463(2002).
 RL Nature 417:459-463(2002).
 DR EMBL; AE012337; AAM41578.1; -; Genomic_DNA.

DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
 DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
 DR InterPro; IPR002508; Amidase_3_hydro.
 DR Pfam; PF01520; Amidase_3_1.
 DR SMART; SM00646; Ami_3_1.
 KW Complete proteome.
 SQ SEQUENCE 563 AA; 57270 MW; 5860B59189C35941 CRC64;

Query Match 25.0%; Score 67; DB 2; Length 563;
 Best Local Similarity 46.9%; Pred. No. 31;
 Matches 15; Conservative 3; Mismatches 12; Indels 2; Gaps 1;

Qy 5 PRPGAPCRAGACTIVLTSGRRQTLSHGSSSPA 36

Db 229 PRPGTVAAPGAP--ATTGGTRTTIATGTPSPA 258

RESULT 12

Q9M4X9_CHLRE PRELIMINARY; PRT; 1787 AA.
 ID Q9M4X9_CHLRE PRELIMINARY; PRT; 1787 AA.
 AC Q9M4X9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Flagellar autotomy protein Falp.
 GN Name=Fal;
 OS Chlamydomonas reinhardtii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
 OC Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
 OX NCBI_TaxID=3055;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=20267878; PubMed=10806107;
 RA Finst R.J., Kim P.J., Griffis E.R., Quarby L.M.;
 RA "Falp is a 171 kDa protein essential for axonemal microtubule severing in Chlamydomonas";
 RL J. Cell Sci. 113:1963-1971(2000).
 DR EMBL; AF246990; AAF66419.1; -; mRNA.
 DR HSP; P37595; 1JN9.
 DR GO; GO:0019861; C:flagellum; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR002197; HTH_Fis.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR Pfam; PF00560; LRR_1; 1.
 DR PRINTS; PR01590; HTHFIS.
 KW Flagellum; leucine-rich repeat; Repeat.
 SQ SEQUENCE 1787 AA; 171585 MW; 452A742595EE14DC1 CRC64;

Query Match 24.8%; Score 66.5; DB 2; Length 1787;
 Best Local Similarity 38.6%; Pred. No. 1.3e+02;
 Matches 22; Conservative 4; Mismatches 14; Indels 17; Gaps 3;

Qy 2 AAG----PRGAPCRAGACTIVL-----TSGRRQTL-----SHGSSSPARATLG 41
 Db 230 AAGGGSAPRPGTAGQAGATAVLGGSLMPGPGQQSLPLSGPSASSSQPGAVTLG 286

RESULT 13

Q743X3_MYCPA PRELIMINARY; PRT; 272 AA.
 ID Q743X3_MYCPA PRELIMINARY; PRT; 272 AA.
 AC Q743X3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocuNames=MAP0468;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OX NCBI_TaxID=1770;

```

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF017228; AAS02785.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00561; Abhydrolase_1; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 272 AA; 30094 MW; 07DB5A382CC8B749 CRC64;

Query Match 24.6%; Score 66; DB 2; Length 272;
Best Local Similarity 40.4%; Pred. No. 18;
Matches 19; Conservative 2; Mismatches 8; Indels 18; Gaps 3;

QY 4 GPRGPACR-----AGAPTIVLTSGRRQTLHGSSSPARATLGKPLV 46
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 7 GPRGPACRCTQDGMPSVLT-----RHUGH-----GEPLV 39

RESULT 14
QY Q8IV50 HUMAN PRELIMINARY; PRT; 215 AA.
AC Q8IV50_
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Hypothetical protein MGC35274 (Putative peptidoglycan binding domain
DE protein-like protein).
GN Name=MGC35274;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Ding P., Han W., Wang Y., Rui M., Chen Y., Wang L., Ma D.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033515; AAH33515.1; -; mRNA.
DR EMBL; AY125955; AAM94507.1; -; mRNA.
DR Ensembl; ENSG00000140280; Homo sapiens.

DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR InterPro; IPR002482; LysM.
DR Pfam; PF01476; LysM; 1.
DR SMART; SMO0257; LysM; 1.
KW Hypothetical protein.
SQ SEQUENCE 215 AA; 23463 MW; A87E666478E903AC CRC64;

Query Match 24.1%; Score 64.5; DB 2; Length 215;
Best Local Similarity 34.0%; Pred. No. 21;
Matches 17; Conservative 11; Mismatches 13; Indels 9; Gaps 2;

QY 5 PRPGAP-----CRAGAPT-----IVLTSGRRQTLHGSSSPARATLGKPLV 45
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 18 PRPSAPSPPRSRSGSEAEELSRLARTKTRSTGTSASVRAPLGAGVI 67

RESULT 15
QY Q82NBL STRAW PRELIMINARY; PRT; 596 AA.
AC Q82NBL_
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Putative transcriptional regulator.
DE OrderedLocusNames=SAV1392;
GN Streptomyces avermitilis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.2114333198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC69102.1; -; Genomic DNA.
GO; GO:0003700; F:transcription factor activity; IEA.
GO; GO:000160; P:two-component signal transduction system (p. .; IEA.
DR InterPro; IPR003018; GAF.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR002197; HTH_Fis.
DR Pfam; PF01590; GAF; 1.
DR Pfam; PF02954; HTH_8; 1.
DR Pfam; PF00158; Sigma54_activat; 1.
DR PRINTS; PR01590; HTHFIS.
DR TIGRFAMs; TIGR01199; HTH_fis; 1.
DR PROSITE; PS00045; SIGMA54_INTERACT_4; 1.
KW Complete proteome.
SQ SEQUENCE 596 AA; 65277 MW; ADEE4EAAE22DF77D CRC64;

Query Match 24.1%; Score 64.5; DB 2; Length 596;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 18; Conservative 3; Mismatches 10; Indels 5; Gaps 1;

QY 11 CR-----AGAPTIVLTSGRRQTLHGSSSPARATLG 41
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 288 CRDSDAGAVFRVRLIGRLQPTSGASPARSLG 323

```



```

FT NON TER 1
SQ SEQUENCE 377 AA; 41650 MW; 0EBDFCF524A4CD6 CRC64;

Query Match
Best Local Similarity 40.5%; 2; Mismatches 20; Indels 3; Gaps 1;
Matches 17; Conservative

QY 5 PR---PGAPCRAGATVILTSRRQTLHGSSSPARATLQK 43
DB 290 PRPGVGTGPMGASGPGYPLRGGRAPSPGYPOOSPYATGGRK 331

RESULT 20
QBQZ77_MOUSE QBQZ77_MOUSE PRELIMINARY; PRT; 562 AA.
AC QBQZ77;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
DE enriched library, clone:B230338M15 product:DEAD/H (ASP-Glu-Ala-
DE Asp/His) box polypeptide 26, full insert sequence. (Fragment).
GN Name=Ddx26;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN RIKENOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN RIKENOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Kleeschmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN RIKENOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN RIKENOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayase N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

```

"Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).

NUCLEOTIDE SEQUENCE.

STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina; MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600; Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.; "RIKEN integrated sequence analysis (RISA) system-384-format" RIKEN integrated sequence analysis (RISA) system-384-format Genome Res. 10:1757-1771(2000).

NUCLEOTIDE SEQUENCE.

STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina; Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki R., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases. DR EMBL; AK046055; BAC32584.1; -; mRNA. DR MGI; MGI:1202397; Ddx26. DR GO; GO:0005615; C:extracellular space; TAS. DR InterPro; IPR000629; DEAD_box. DR PROSITE; PS00039; DEAD_ATP_HELICASE; UNKNOWN_1. FT NON TER 1. SQ SEQUENCE 562 AA; 64142 MW; 233C3BCERAS1D4A7 CRC64;

Query Match 23.7%; Score 63.5; DB 2; Length 562; Best Local Similarity 37.7%; Pred. No. 81; Matches 23; Conservative 2; Mismatches 21; Indels 15; Gaps 4;

QY 3 AGP-----RPGAPCRAGAP-----TIVLTSGRRQT---LSH--GSSSPARATLQKPLVLN 47
 DB 324 AGPQKHKRPGEPSMQGIPKRRRCASPLLRGRQSPAVNSHIGKGPPAPMTQAPETN 383
 QY 48 D 48
 DB 384 D 384

RESULT 21
 QBPD77_MOUSE QBPD77_MOUSE PRELIMINARY; PRT; 874 AA.
 ID QBPD77; AC QBPD77;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Ddx26 protein.
 GN Name=Ddx26;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN RIKENOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,

RA Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Director MGC Project;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058637; AAH58637.1; -; mRNA.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000629; DEAD box.
DR InterPro; IPR002035; VWF_A.
DR PRINTS; PR00453; VWFADOMAIN.
DR PROSITE; PS00039; DEAD ATP HELICASE; UNKNOWN_1.
DR PROSITE; PS50234; VWF_A1.
SQ SEQUENCE 874 AA; 98634 MW; BA168BF36975A235B CRC64;

Query Match 23.7%; Score 63.5; DB 2; Length 874;
Best Local Similarity 37.7%; Pred. No. 1.3e+02;
Matches 23; Conservative 2; Mismatches 21; Indels 15; Gaps 4;
4;
Oy 3 AGP-----RCPACRCAGAP-----TIVLTSGRRGT-----LSH--GSSSPARATLTKPLVLN 47
Db 536 AGPQNKHKRGPSPSMQGIKRRRCASPLLRGRSPVNSHIGKGPPAPWTAQPEATN 695
Oy 48 D 48
Db 696 D 696

RESULT 22
Q8NOY6 HUMAN
ID Q8NOY6 HUMAN PRELIMINARY; PRT; 128 AA.
AC Q8NOY6
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein FLJ35271 (Hypothetical protein FLJ37078).
GN Name=FLJ37078;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Cerebellum, and Prostate;
RA PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Ohtsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Katsumi N., Sato K., Tanikawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,

RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani K., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RG NIH MGC Project;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK092590; BAC03922.1; -; mRNA.
DR EMBL; AK094397; BAC04346.1; -; mRNA.
DR EMBL; BC063788; AAH63788.1; -; mRNA.
DR Ensembl; ENSG0000017679; Homo sapiens.
KW Hypothetical protein.
SQ SEQUENCE 128 AA; 13058 MW; BC78007011570EB5 CRC64;

Query Match 23.5%; Score 63; DB 2; Length 128;
Best Local Similarity 35.6%; Pred. No. 17;
Matches 16; Conservative 4; Mismatches 23; Indels 2; Gaps 1;
1;
Oy 4 GPRPGAPCRAGATTIVLTSGRRQTLSHGSSSPARATLTKPLVLND 48
Db 2 GPAAGHAAPRGPPVPATTAG--AAALRAGASEPRTQLGAPLALGE 44

RESULT 23
Q6K3G1 ORYSA
ID Q6K3G1 ORYSA PRELIMINARY; PRT; 171 AA.
AC Q6K3G1
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP	[3]
RP	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Placenta;
RC	Director MGC Project;
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; MI9720; AAA59878.1; -; Genomic DNA.
DR	EMBL; BC011864; AAH11864.1; -; mRNA.
DR	TRANSFAC; T02386;
DR	GO; GO:0005634; C:nucleus; NAS.
DR	GO; GO:0003700; F:transcription factor activity; NAS.
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR	InterPro; IPR002418; TF_Myc.
DR	Pfam; PF01056; MYC_N; 1.
DR	PRINTS; PR00044; LEUZIPRMYC.
DR	SEQUENCE 206 AA; 21766 MW; 585C9CD6C9A8EC71 CRC64;
QY	Query Match 23.5%; Score 63; DB 2; Length 206;
Db	Best Local Similarity 33.3%; Pred. No. 29;
	Matches 14; Conservative 1; Mismatches 27; Indels 0; Gaps
	2 AAGPPGAPCPAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
	:
	137 AGNPAPAAPCLGPBKTKOACSGSEPSDSKDLPEPSKRGGP 178
	:
RESULT 25	
OSQPLO HUMAN	
ID OSQPLO HUMAN PRELIMINARY; PRT; 236 AA.	
AC OSQPLO;	
DT 01-FEB-2005 (TReMBLrel. 29, Created)	
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)	
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)	
DE V-myc myelocytomatosis viral oncogene homolog 1, lung carcinoma	
DE derived (Avian).	
GN Name=MYCL1; ORFNames=RPL-118J21.4-002;	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;	
OC Homo.	
OX NCBI_TaxID=9606;	
ON [1]	
RP NUCLEOTIDE SEQUENCE.	
RA Ellington A.;	
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AL033527; CAI19408.1; -; Genomic DNA.	
DR GO; GO:0005634; C:nucleus; IEA.	
DR GO; GO:0003700; F:transcription factor activity; IEA.	
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	
DR InterPro; IPR002418; TF_Myc.	
DR Pfam; PF01056; MYC_N; 1.	
DR PRINTS; PR00044; LEUZIPRMYC.	
DR SEQUENCE 236 AA; 24682 MW; 5059EED93CE18391 CRC64;	
QY Query Match 23.5%; Score 63; DB 2; Length 236;	
Db Best Local Similarity 33.3%; Pred. No. 34;	
	Matches 14; Conservative 1; Mismatches 27; Indels 0; Gaps
	2 AAGPPGAPCPAGAPTIVLTSGRRQTLSHGSSSPARATLGKP 43
	:
	167 AGNPAPAAPCLGPBKTKOACSGSEPSDSKDLPEPSKRGGP 208
	:
RESULT 26	
YBOX2 MOUSE	
ID YBOX2 MOUSE STANDARD; PRT; 360 AA.	

AC Q922C8; Q5NCW8; Q5NCW9; Q922C7;
 DT 10-MAY-2005 (Rel. 47, Created)
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Y-box binding protein 2 (Germ cell specific Y-box binding protein)
 DE (PBG2 homolog).
 GN Name=Ybx2; Synonyms=MSY2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.
 RC STRAIN=CD-1; TISSUE=Testis;
 RX MEDLINE=98455458; PubMed=9780336;
 RA Gu W., Tekur S., Reinbold R., Eppig J.J., Choi Y.-C., Zheng J.Z.,
 RA Murray M.T., Hecht N.B.;
 RT "Mammalian male and female germ cells express a germ cell-specific Y-
 box protein, MSY2.";
 RL Biol. Reprod. 59:1266-1274 (1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC The mouse genome sequencing consortium;
 RL Submitted (DEC-2004) to the ENBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION IN MRNA TRANSLATION, PHOSPHORYLATION, RNA-BINDING, AND
 RP IDENTIFICATION IN A MRNP COMPLEX WITH PABPC1.
 RX PubMed=10076007; DOI=10.1093/nar/27.7.1747;
 RA Herbert T.P., Hecht N.B.;
 RT "The mouse Y-box protein, MSY2, is associated with a kinase on non-
 polyosomal mouse testicular mRNAs.";
 RL Nucleic Acids Res. 27:1747-1753 (1999).
 RN [4]
 RP IDENTIFICATION IN A MRNP COMPLEX WITH CSDA.
 RC TISSUE=Testis;
 RX PubMed=10772793; DOI=10.1006/dbio.2000.9658;
 RA Davies H.G., Giorgini F., Fajardo M.A., Braun R.E.;
 RT "A sequence-specific RNA binding complex expressed in murine germ
 cells contains MSY2 and MSY4.";
 RL Dev. Biol. 221:87-100 (2000).
 RN [5]
 RP PHOSPHORYLATION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RX PubMed=11566752;
 RA Yu J., Hecht N.B., Schultz R.M.;
 RT "Expression of MSY2 in mouse oocytes and preimplantation embryos.";
 RL Biol. Reprod. 65:1260-1270 (2001).
 RN [6]
 RP RNA-BINDING.
 RX PubMed=11564883; DOI=10.1128/MCB.21.20.7010-7019.2001;
 RA Giorgini F., Davies H.G., Braun R.E.;
 RT "MSY2 and MSY4 bind a conserved sequence in the 3' untranslated region
 of protamine 1 mRNA in vitro and in vivo.";
 RL Mol. Cell. Biol. 21:7010-7019 (2001).
 RN [7]
 RP FUNCTION IN MRNA TRANSLATION, AND RNA-BINDING.
 RX PubMed=12297523;
 RA Yu J., Hecht N.B., Schultz R.M.;
 RT "RNA-binding properties and translation repression in vitro by germ
 cell-specific MSY2 protein.";
 RL Biol. Reprod. 67:1093-1098 (2002).
 RN [8]
 RP FUNCTION IN MATERNAL MRNA CYTOPLASMIC RETENTION, RNA-BINDING, AND
 RP MUTAGENESIS OF TYR-109 AND PHE-111.
 RX PubMed=12648488; DOI=10.1016/S0012-1606(02)00094-5;
 RA Yu J., Hecht N.B., Schultz R.M.;
 RT "Requirement for RNA-binding activity of MSY2 for cytoplasmic
 localization and retention in mouse oocytes.";
 RL Dev. Biol. 255:249-262 (2003).
 RN [9]
 RP FUNCTION IN MRNA STABILITY.
 RX PubMed=15031116; DOI=10.1016/j.ydbio.2003.12.020;
 RA Yu J., Deng M., Medvedev S., Yang J., Hecht N.B., Schultz R.M.;

RT "Transgenic RNAi-mediated reduction of MSY2 in mouse oocytes results
 in reduced fertility.";
 RL Dev. Biol. 268:195-206 (2004).
 RN [10]
 RP FUNCTION IN CYTOPLASMIC MRNA STORAGE, RNA-BINDING, AND DNA-BINDING.
 RX PubMed=15665108; DOI=10.1073/pnas.040485102;
 RA Yang J., Medvedev S., Reddi P.P., Schultz R.M., Hecht N.B.;
 RT "The DNA/RNA-binding protein MSY2 marks specific transcripts for
 cytoplasmic storage in mouse male germ cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:1513-1518 (2005).
 CC -1- FUNCTION: Major constituent of messenger ribonucleoprotein
 CC particles (MRNPs). Involved in the regulation of the stability
 CC and/or translation of germ cell mRNAs. Binds to Y-box consensus
 CC promoter element. Binds to full length mRNA with high affinity in
 CC a sequence-independent manner. Binds to short RNA sequences
 CC containing the consensus site 5'-UCCAUC-3' with low affinity and
 CC limited sequence specificity. Its binding with maternal mRNAs is
 CC necessary for its cytoplasmic retention. May mark specific mRNAs
 CC (those transcribed from Y-box promoters) in the nucleus for
 CC cytoplasmic storage, thereby linking transcription and mRNA
 CC storage/translational delay.
 CC -1- SUBUNIT: Found in a mRNP complex with PABPC1 and CSDA.
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=MSY2;
 CC IsoId=Q922C8-1; Sequence=Displayed;
 CC Name=2; Synonyms=MSY2a;
 CC IsoId=Q922C8-2; Sequence=VSP_013052, VSP_013053;
 CC Note=In Ref.4 isoform 2 not detected in oocytes;
 CC -1- TISSUE SPECIFICITY: Expressed in meiotic and postmeiotic male germ
 CC cells and oocytes; poorly expressed in two cell stage embryos (at
 CC protein level). Not detected in preimplantation embryos.
 CC -1- PTM: Phosphorylated during oocyte maturation and dephosphorylated
 CC following egg activation. Phosphorylation in vitro by a kinase
 CC activity associated with testicular mRNPs. Dephosphorylation leads
 CC to a decrease in its affinity to bind RNA in vitro.
 CC -1- MISCELLANEOUS: Knockout mice for this gene exhibited reduced level
 CC of mRNA and protein synthesis in fully grown oocytes and decrease
 CC fertility.
 CC -1- SIMILARITY: Contains 1 CSD (cold-shock) domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF073954; AAC98673.1; -; mRNA.
 CC EMBL; AF073955; AAC98674.1; -; mRNA.
 CC EMBL; AL596185; CAI35155.1; -; Genomic DNA.
 CC EMBL; AL596185; CAI35156.1; -; Genomic DNA.
 CC HSPSP; P41016; 1HZA.
 CC SMR; Q922C8; 89-166.
 CC Ensembl; ENSMUSG00000018554; Mus musculus.
 CC MGI; MGI:1096372; Ybx2.
 CC GO; GO:0005737; Cytoplasm; IDA.
 CC GO; GO:0003729; F:mRNA binding; IDA.
 CC GO; GO:0045182; F:translation regulator activity; IDA.
 CC InterPro; IPR011129; CSP.
 CC InterPro; IPR020259; CSP_DNA_bd.
 CC InterPro; IPR012340; OB_NA_bd_sub.
 CC Pfam; PF00313; CSD; 1.
 CC PRINTS; PR00050; COLDSHOCK.
 CC ProDom; PD000621; Cold shock; 1.
 CC SMART; SM00357; CSP; 1.
 CC PROSITE; PS00352; COLD_SHOCK; 1.
 CC Alternative splicing; DNA-binding; Nuclear protein; Phosphorylation;
 CC RNA-binding.
 CC DOMAIN 95 165 CSD.
 CC REGION 89 171 Required for cytoplasmic retention.
 CC REGION 219 360 Required for mRNA-binding.
 CC COMPTAS 41 57 Gly-rich.

```

FT COMBIAS 174 241 Arg-rich.
FT VARSPLIC 1 78 Missing (in isoform 2).
FT /FTid=VSP_013052.
FT PAPPARQADKPV -> MSRRAGQAGSAKA (in
FT isoform 2).
FT VARSPLIC 79 91
FT MUTAGEN 109 109 Y->A: Abolishes cytoplasmic retention and
FT mRNA-binding affinity; when associated
FT with A-111.
FT MUTAGEN 111 111 F->A: Abolishes cytoplasmic retention and
FT mRNA-binding affinity; when associated
FT with A-109.
FT CONFLICT 60 61 LH -> S (in Ref. 2; CAI35155).
FT CONFLICT 158 158 R -> E (in Ref. 2).
SQ SEQUENCE 360 AA; 38271 MW; DFCAE7F936731BF CRC64;

Query Match 23.5%; Score 63; DB 1; Length 360;
Best Local Similarity 35.6%; Pred. No. 56;
Matches 16; Conservative 8; Mismatches 17; Indels 4; Gaps 2;

Qy 2 AAGPRGAPCRAGAPTIVLTSGRRQTLSHGSS-SPARATLQKPLV 45
Db 51 ASGPAAGTFLHAPGPR---TPGNOATAASGTPAPPARSQADKPV 92

RESULT 27
Q7WNE3 BORBR PRELIMINARY; PRT; 373 AA.
AC Q7WNE3;
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
RA Parthill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RL Nat. Genet. 35:32-40(2003).
DR EMBL; EX640440; CAE31595.1; -; Genomic_DNA.
DR HSSP; P77407; 1PQY.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB.BAIF.
DR Pfam; PF02515; Coa_transf_3; 1.
KW Complete proteome.
SQ SEQUENCE 373 AA; 39854 MW; 61079E192969E806 CRC64;

Query Match 23.5%; Score 63; DB 2; Length 373;
Best Local Similarity 48.0%; Pred. No. 58;
Matches 12; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 GPRPGAPCRAGAPTIVLTSGRRQTL 28
Db 165 GPADGGPCRCVGPVIVDLVTGNAAAL 189

RESULT 28

```

```

O4QG78 LEIMA
ID Q4QG78 LEIMA PRELIMINARY; PRT; 2063 AA.
AC Q4QG78;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Lmjf13.0760;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Apostolou Z., Bauser C., Beck A.,
RA Blanchettin G., Borzym K., Bothe G., Bruchi C., Ciaroni L.,
RA Diesterhoef A., Fuchs M., Gabel C., Goffeau A., Hilbert H.,
RA Klages S., Kube M., Marra M., Masuy D., Mueller-Auer S., Pohl T.,
RA Purnelle B., Reinhardt R., Rieger M., Robben J., Schaefer M.,
RA Tosato V., Volckaert G., Wambutt R., Wedler H., Zimmermann M.,
RA Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005252; CAJ02713.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 2063 AA; 218311 MW; 78E1446A6E0976C0 CRC64;

Query Match 23.5%; Score 63; DB 2; Length 2063;
Best Local Similarity 31.0%; Pred. No. 4e+02;
Matches 18; Conservative 8; Mismatches 12; Indels 20; Gaps 2;

Qy 3 AGPRPGAPCRAG-----APTIVLTSGRRQTLSHGSS-----SPARATL 40
Db 1300 AGATPGSPSRGATLPKSRSPSAVTSRRPYSSSSSSSSSTSSRTLTSPSKASL 1357

RESULT 29
O84ZQ6 ORYSA
ID O84ZQ6 ORYSA PRELIMINARY; PRT; 150 AA.
AC O84ZQ6;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein OJ1372_D12.117.
GN Names=OJ1372_D12.117;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ1372_D12.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003827; BAC57658.1; -; Genomic_DNA.
DR Gramene; O84ZQ6; -
KW Hypothetical protein.
SQ SEQUENCE 150 AA; 15791 MW; F5A09E6767446CFB CRC64;

Query Match 23.3%; Score 62.5; DB 2; Length 150;
Best Local Similarity 43.1%; Pred. No. 24;
Matches 22; Conservative 2; Mismatches 18; Indels 9; Gaps 2;

Qy 2 AAGPRPGAPCRAGAPTIVLTSGRRQTLSHGSSPA-----RATLGRP 43
Db 20 AARPLANHPNIRQAPACISGRSSSLSPSSSPASPLLSLRSSGKP 70

RESULT 30
O529G0_ORYSA

```

```

ID Q529G0_ORYSA PRELIMINARY; PRT; 208 AA.
AC Q529G0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Ferredoxin oxidoreductase a-subunit.
DE Hypothetical protein P0655A07.32.
GN Name=P0655A07.32;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0655A07."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003634; BAD53647.1; -; Genomic_DNA.
DK Hypothetical protein.
SQ SEQUENCE 208 AA; 20992 MW; 455913B4183221D7 CRC64;

Query Match 23.3%; Score 62.5; DB 2; Length 208;
Best Local Similarity 45.08; Pred. No. 34;
Matches 18; Conservative 1; Mismatches 16; Indels 5; Gaps 2;

QY 5 PRPGAPCRA-GAPTIVTSGRRQTLTSHGSSSPARATLGRP 43
| | | | | | | | | | | | | | | | | | | | | |
Db 58 PPQAPCGRRRLPTTL-----RSAPSSSSMPPPTATLDP 93

RESULT 31
Q8XQ41_RALSO
ID Q8XQ41_RALSO PRELIMINARY; PRT; 556 AA.
AC Q8XQ41;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE PROBABLE ACTIVATION/SECRETION SIGNAL PEPTIDE PROTEIN.
GN OrderedLocuNames=RSpl445; ORFNames=RS03098;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GMT1000;
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siquier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weltschbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646084; CAD18596.1; -; Genomic_DNA.
KW Complete proteome; Plasmid.
SQ SEQUENCE 556 AA; 58835 MW; 5C27CF4AAED0EC72 CRC64;

Query Match 23.3%; Score 62.5; DB 2; Length 556;
Best Local Similarity 37.5%; Pred. No. 1e+02;
Matches 18; Conservative 7; Mismatches 22; Indels 1; Gaps 1;

QY 1 MAAGPRPG-APCRAGAPTIVTSGRRQTLTSHGSSSPARATLGRP 47
| | | | | | | | | | | | | | | | | | | | | |
Db 202 LAAGARPGQTLVIGSRAAPLTSGRLADNYGSLYTGRTVGGTLNVN 249

RESULT 32
Q925X4_FRASE
ID Q925X4_FRASE PRELIMINARY; PRT; 671 AA.

```

```

AC Q925X4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ferredoxin oxidoreductase a-subunit.
OS Frankia sp. (strain EuIK1).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Frankineae; Frankiaceae; Frankia.
OX NCBI_TaxID=47227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EuIK1;
RA Chung-Sun A., Ji-Tae K., Won-Jin K., Won-Young Y.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF119361; AAD17273.1; -; Genomic DNA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002880; POR_N.
DR Pfam; PF01855; POR_N; 1.
SQ SEQUENCE 671 AA; 71283 MW; 3EAF14BBB36C598 CRC64;

Query Match 23.3%; Score 62.5; DB 2; Length 671;
Best Local Similarity 38.3%; Pred. No. 1.3e+02;
Matches 18; Conservative 5; Mismatches 17; Indels 7; Gaps 2;

QY 1 MAAGPRPG-APCRAGAPTIVTSGRRQTLTSHGSSSPARATLKG 42
| | | | | | | | | | | | | | | | | | | | | |
Db 540 LPARPRPGAGHRPGAPALQPAAGEPRGRAVVLRHGAG--AEMNLGQ 584

RESULT 33
Q5ZKM6_CHICK
ID Q5ZKM6_CHICK PRELIMINARY; PRT; 771 AA.
AC Q5ZKM6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN ORFNames=RCJMB04_10a9;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kutter S., Biagodatki A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis."
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ720058; CAG31717.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 771 AA; 88312 MW; B26F8692352887F CRC64;

Query Match 23.3%; Score 62.5; DB 2; Length 771;
Best Local Similarity 42.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 1; Mismatches 17; Indels 5; Gaps 1;

QY 2 AAGPRPG-APCRAGAPTIVTSGRRQTLTSHGSSSPARATLGG 41
| | | | | | | | | | | | | | | | | | | | | |
Db 3 AVGGPGGRRRRRGAPTILAAQ-----LRHQLSGRGLTRG 37

RESULT 34
O93392_COTCO
ID O93392_COTCO PRELIMINARY; PRT; 215 AA.
AC O93392;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

```
DE Pro-alpha2(I) collagen (Fragment).
OS Coturnix coturnix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99119493; PubMed=9918878; DOI=10.1006/viro.1998.9499;
RA Oberst C., Hartl M., Weiskirchen R., Bister K.;
RT "Conditional cell transformation by doxycycline-controlled expression
of the MC29 v-myc allele."
RL Virology 253:193-207(1999).
DR EMBL: AF077329; AD12181.1; -; mRNA.
DR GO: GO:0005737; Cytoplasm; IEA.
DR GO: GO:0005515; F:protein binding; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0006817; P:phosphate transport; IEA.
DR InterPro: IPR008161; Clg helix.
DR InterPro: IPR008160; Collagen.
DR Pfam: PF01391; Collagen; 3.
DR ProDom: PD000007; Clg_helix; 3.
KW Collagen.
FT NON TER
SQ SEQUENCE 215 AA; 20194 MW; CDB607AE07D03126 CRC64;

Query Match 23.1%; Score 62; DB 2; Length 215;
Best Local Similarity 37.5%; Pred. No. 41;
Matches 21; Conservative 4; Mismatches 17; Indels 14; Gaps 3;

QY 3 AGPP-----PGAPCRAGAPTIVLTGRRQTLHGSSSPARATLGKPLVNDYS 50
DB 158 AGPPGPPGPPGKGRNGLPPIGPAGVRG--SHGSQGPA----GPPVPLTWS 207

RESULT 35
Q6ZM28 HUMAN
ID Q6ZM28_HUMAN PRELIMINARY; PRT; 259 AA.
AC Q6ZM28;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein FLJ16555.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Splice;
RA Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
RA Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A., Iehi S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Takahashi-Fujii K., Oshima A., Sugiyama A., Kawakami B., Kanehori K.,
RA Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL: AK131430; BAD19577.1; -; mRNA.
DR HSSP: F15822; IBB0.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0046872; F:metal ion binding; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR InterPro: IPR007087; ZnF_C2H2.
DR Pfam: PF00096; zf-C2H2; 2.
DR SMART: SM00355; ZnF_C2H2; 4.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 2.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 259 AA; 29243 MW; 8C8026415B7559C4 CRC64;
```

```
Query Match 23.1%; Score 62; DB 2; Length 415;
Best Local Similarity 37.8%; Pred. No. 86;
Matches 14; Conservative 3; Mismatches 20; Indels 0; Gaps 0;

QY 5 PRGAPCRAGAPTIVLTGRRQTLHGSSSPARATLG 41
DB 226 PSPTPCRSPADPTDTASPPNVSPSSSPASPATG 262

RESULT 37
Q7SG58 NEUCR
ID Q7SG58_NEUCR PRELIMINARY; PRT; 449 AA.
AC Q7SG58;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein.
DE Hypothetical protein.
GN Name=NCU07482.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
```

```
Query Match 23.1%; Score 62; DB 2; Length 259;
Best Local Similarity 30.4%; Pred. No. 50;
Matches 21; Conservative 5; Mismatches 19; Indels 24; Gaps 3;

QY 5 PRGAPCRAGAPTIVLTGRR-----QTL-----SHGSSSPARATL----- 40
DB 135 PAPCPAGPAPLLVVGSRNLLSPLSVNASQALQTVLSAAHGSSEPNLAKALAFN 194

QY 41 GKPLVLDY 49
DB 195 GSPLRFDKY 203

RESULT 36
Q6A047 MOUSE
ID Q6A047_MOUSE PRELIMINARY; PRT; 415 AA.
AC Q6A047;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE MKIAA0561 protein (Fragment).
GN Name=Mast3; Synonyms=mkIAA0561;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Adult thymus;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
Randomly Sampled from Size-Fractionated Libraries."
RL DNA Res. 11:205-218(2004).
CC -1- SIMILARITY: Contains 1 PDZ (DHR) domain.
DR EMBL: AK172971; BAD32249.1; -; mRNA.
DR MGI: MGI:2683541; Mast3.
DR GO: GO:0005515; F:protein binding; IEA.
DR InterPro: IPR001478; PDZ.
DR Pfam: PF00595; PDZ; 1.
DR SMART: SM00228; PDZ; 1.
DR PROSITE: PS0106; PDZ; 1.
FT NON TER
SQ SEQUENCE 415 AA; 43869 MW; FF7A39D8ACFA96D0 CRC64;
```


DT 10-MAY-2005 (TREMELrel. 30, Created)
DT 10-MAY-2005 (TREMELrel. 30, Last sequence update)
DE 10-MAY-2005 (TREMELrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNE05190;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
[1]
RP NUCLEOTIDE SEQUENCE.
PC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
PC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung K.J., Lenggeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Mathewson C.A., Mitchell T.G., Perte M., Riggs F.R., Salzberg S.L.,
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
Cryptococcus neoformans.";
RL Science 307:1321-1324 (2005).
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL: AE017345; AAW43739.1; -: Genomic_DNA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; RVT 1; 1.
DR PROSITE: PS50878; RT_POL; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1059 AA; 116488 MW; 79F19996630E066B CRC64;

Query Match 23.1%; Score 62; DB 2; Length 1059;
Best Local Similarity 29.7%; Pred. No. 2.5e+02;
Matches 19; Conservative 8; Mismatches 15; Indels 22; Gaps 3;

QY 1 MAAGPRPGAPCAG-----APTIVLTSGRRQTLHSGSSSPARA---TLGKP 43
49 LVACPRGTCGAGTGLSRHQRYCGLTAPRI-----RRNRVGNSTNTPRPPSNTAASP 103

QY 44 LVLN 47
104 IVLS 107

RESULT 41
Q55885 CRYNE PRELIMINARY; PRT; 1059 AA.
AC Q55885;
DT 13-SEP-2005 (TREMELrel. 31, Created)
DT 13-SEP-2005 (TREMELrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMELrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNEE2370;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;

RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -!- SIMILARITY: Contains 1 reverse transcriptase domain.
DR EMBL: AAEY01000025; EAL20632.1; -: Genomic_DNA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO: GO:006278; P:RNA-dependent DNA replication; IEA.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; RVT 1; 1.
DR PROSITE: PS50878; RT_POL; 1.
KW Hypothetical protein.
SQ SEQUENCE 1059 AA; 116568 MW; B52AE531A2E90F1B CRC64;

Query Match 23.1%; Score 62; DB 2; Length 1059;
Best Local Similarity 29.7%; Pred. No. 2.5e+02;
Matches 19; Conservative 8; Mismatches 15; Indels 22; Gaps 3;

QY 1 MAAGPRPGAPCAG-----APTIVLTSGRRQTLHSGSSSPARA---TLGKP 43
49 LVACPRGTCGAGTGLSRHQRYCGLTAPRI-----RRNRVGNSTNTPRPPSNTAASP 103

QY 44 LVLN 47
104 IVLS 107

RESULT 42
RIMB2 CHICK STANDARD; PRT; 1325 AA.
ID RIMB2 CHICK
AC Q80FX1;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 29-MAY-2005 (Rel. 47, Last annotation update)
DE RIM binding protein 2 (RIM-BP2).
GN Name=RIMBP2; Synonyms=RB2;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE. FUNCTION, SUBCELLULAR LOCATION, TISSUE
RP SPECIFICITY, INTERACTIONS WITH RIMS1, RIMS2, CACNA1D AND CACNA1B, AND
RP FUNCTION OF SH3 DOMAIN.
RX MEDLINE=21984561; PubMed=11988172; DOI=10.1016/S0896-6273(02)00667-0;
RA Hibino H., Pironkova R., Onumure O., Volododskaya M., Hudspeth A.J.,
RA Lesage F.;
RT "RIM binding proteins (RBPs) couple Rab3-interacting molecules (RIMs)
RT to voltage-gated Ca(2+) channels.";
RL Neuron 34:411-423 (2002).
CC -!- FUNCTION: Plays a role in the synaptic transmission as
CC bifunctional linker that interacts simultaneously with RIMS1,
CC RIMS2, CACNA1D and CACNA1B.
CC -!- SUBUNIT: Interacts with RIMS1, RIMS2, CACNA1D and CACNA1B, and
CC potentially with other Ca(2+) channel alpha-1 isoforms.
CC -!- SUBCELLULAR LOCATION: Synaptic plasma membrane.
CC -!- TISSUE SPECIFICITY: Brain, cochlea and retina.
CC -!- DOMAIN: The SH3 domains mediate binding to a proline-rich motif in
CC RIMS1, RIMS2, CACNA1D and CACNA1B.
CC -!- SIMILARITY: Contains 3 fibronectin type-III domains.
CC -!- SIMILARITY: Contains 3 SH3 domains.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

EMBL; AY072908; AAL67995.1; -, mRNA.
DR HSP; P29354; IAZE.
DR SMR; Q8QFX1; 181-261.
DR Ensembl; ENSGALG00000002579; Gallus gallus.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR011511; SH3_2.
DR Pfam; PF00041; FN3; 3.
DR Pfam; PF07653; SH3_2; 3.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 3.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00326; SH3; 3.
DR PROSITE; PS00853; FN3; 3.
DR PROSITE; PS00002; SH3; 3.
KW Membrane; Repeat; SH3 domain.
FT DOMAIN 186 253 SH3 1.
FT DOMAIN 315 395 Fibronectin type-III 1.
FT DOMAIN 409 481 Fibronectin type-III 2.
FT DOMAIN 505 591 Fibronectin type-III 3.
FT DOMAIN 1121 1189 SH3 2.
FT DOMAIN 1225 1292 SH3 3.
SQ SEQUENCE 1325 AA; 147538 MW; 1DB8C50D4CD41B5B CRC64;
Query Match 23.1%; Score 62; DB 1; Length 1325;
Best Local Similarity 34.5%; Pred. No. 3.2e+02;
Matches 20; Conservative 7; Mismatches 13; Indels 18; Gaps 3;
QY 2 AAGPRP-----GAPCRAGAPTIVLTSGRRQTLTGKPLV 46
DQ 1040 AAGPLAKSHRDARDSDLLGNPSSAGRPVEHARRS--SHGSAVQRS---RPMIV 1092
DB
RESULT 43
Q4SE53 TETNG
ID Q4SE53 TETNG PRELIMINARY; PRT; 1553 AA.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF14625, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0019708001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozuof-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Aichmond V., Jubin C., Castellino V., Poulain J., De Berardinis V.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., Gouzy J.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Croliis H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; CAAB01014625; CAG01079.1; -, Genomic DNA.
FT NON TER 1
SQ SEQUENCE 1553 AA; 166611 MW; 47624318FEA7B01C CRC64;
Query Match 23.1%; Score 62; DB 2; Length 1553;
Best Local Similarity 35.7%; Pred. No. 3.8e+02;
Matches 20; Conservative 6; Mismatches 24; Indels 6; Gaps 3;
QY 3 AGRPRGAPCRAG-AP-TIVLTSGRRQTLTGKPLV 52
DQ 476 SGPEPRCCSTGSGAPGVRSGEGEATRPDGGGVSSSPASSAQAGPAALPEVQVD 531
DB
RESULT 44
Q9N2G7 RABIT
ID Q9N2G7 RABIT PRELIMINARY; PRT; 588 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Eukaryotic polypeptide chain release factor 3 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Karamysheva Z.N., Karamyshev A.L., Ito K., Yokogawa T., Nishikawa K.,
RA Nakamura Y., Matsufuji S.;
RT "Antizyme frameshifting as a functional probe of eukaryotic
RT translational termination.";
RL Nucleic Acids Res. 31:5949-5956(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Karamysheva Z., Matsufuji S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB035256; BAA92160.1; -, mRNA.
DR HSP; P02990; IEFU.
DR GO; GO:0005622; C:intracellular; ISS.
DR GO; GO:0003924; F:GTPase activity; ISS.
DR GO; GO:0005151; F:protein binding; ISS.
DR GO; GO:0003747; F:translation release factor activity; ISS.
DR GO; GO:0000082; P:G1/S transition of mitotic cell cycle; ISS.
DR GO; GO:0000184; P:mRNA catabolism, nonsense-mediated decay; ISS.
DR InterPro; IPR009818; Ataxin-2 C.
DR InterPro; IPR004160; EFTU Cterm.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR000795; ProSyn_GTPbind.
DR Pfam; PF00009; GTP_EFTU_1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR Pfam; PF07145; PAM2; 1.
DR PRINTS; PR00315; ELONGATNPT.
DR PROSITE; PS00301; EFATOR GTP; UNKNOWN 1.
KW GTP-binding; Nucleotide-binding; Protein biosynthesis.
FT NON TER 1
SQ SEQUENCE 588 AA; 64476 MW; 082A3BCA575777AE CRC64;
Query Match 22.9%; Score 61.5; DB 2; Length 588;
Best Local Similarity 40.4%; Pred. No. 1.5e+02;
Matches 19; Conservative 5; Mismatches 22; Indels 1; Gaps 1;
QY 2 AAGPRGAPCRAGAPTIVLTSGRRQTLTGKPLV 47
DQ 53 AAGNHGAGSAGPSAPVSSQEQSLCEGSIASVMELSEPVVEN 99
DB
RESULT 45
Q69ZH8 MOUSE
ID Q69ZH8 MOUSE PRELIMINARY; PRT; 1178 AA.
AC Q69ZH8
DT 25-OCT-2004 (TrEMBLrel. 28, Created)


```

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
GN MKIAA1509 protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AK173188; BAD32466.1; -; -.
DR MGI; MGI:1915589; 0610010D24Rik.
FT NON_TER 1
SQ SEQUENCE 1178 AA; 130942 MW; 0E44EF5C79F675BA CRC64;

Query Match 22.9%; Score 61.5; DB 2; Length 1178;
Best Local Similarity 27.8%; Pred. No. 3.2e+02;
Matches 20; Conservative 9; Mismatches 18; Indels 25; Gaps 3;

QY 2 AAGP---RGAFCRAGAPTIVLTSGRRQTLSHG-----SSSPARAT 39
DB 976 ASGPEACRPESQKPGGHE---AGARETSTHSLQGSILARERTPIVGKADSPSGQGT 1032
QY 40 LGKPLVLNDYSL 51
DB 1033 RGRPLDTRRFSL 1044

```

Search completed: December 27, 2005, 23:04:18
Job time : 230 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 22:57:02 ; Search time 46 Seconds
(without alignments)
93.459 Million cell updates/sec

Title: US-09-819-144A-2

Perfect score: 268
Sequence: 1 MAAGPRGAPCRAGAPTIVL.....SSPARATLGKPLVNDYSLN 52

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/1aa/5 COMB.pap.*
- 2: /cgn2_6/ptodata/1/1aa/6 COMB.pap.*
- 3: /cgn2_6/ptodata/1/1aa/H COMB.pap.*
- 4: /cgn2_6/ptodata/1/1aa/PTUS COMB.pap.*
- 5: /cgn2_6/ptodata/1/1aa/RE COMB.pap.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	268	100.0	52	2	US-09-042-225-2	Sequence 2, Appli
2	74	27.6	16	2	US-09-042-225-8	Sequence 8, Appli
3	72.5	27.1	173	2	US-09-252-991A-25321	Sequence 25321, A
4	67.5	25.2	137	2	US-09-252-991A-30540	Sequence 30540, A
5	66	24.6	228	2	US-09-252-991A-28381	Sequence 28381, A
6	65	24.3	166	2	US-09-252-991A-16776	Sequence 16776, A
7	65	24.3	205	2	US-09-252-991A-31167	Sequence 31167, A
8	65	24.3	345	2	US-09-252-991A-25768	Sequence 25768, A
9	64.5	24.1	146	2	US-09-252-991A-25930	Sequence 25930, A
10	64.5	24.1	149	2	US-09-252-991A-30866	Sequence 30866, A
11	63.5	23.7	721	2	US-09-252-991A-28293	Sequence 28293, A
12	63	23.5	128	2	US-10-104-047-3112	Sequence 3112, Ap
13	63	23.5	613	2	US-09-252-991A-19955	Sequence 19955, A
14	63	23.5	1129	2	US-09-252-991A-22330	Sequence 22330, A
15	62	23.1	256	2	US-09-252-991A-25404	Sequence 25404, A
16	62	23.1	328	2	US-09-252-991A-19582	Sequence 19582, A
17	62	23.1	863	2	US-09-252-991A-26099	Sequence 26099, A
18	61.5	22.9	146	2	US-09-949-016-8300	Sequence 8300, Ap
19	61.5	22.9	367	2	US-09-252-991A-19910	Sequence 19910, A
20	61.5	22.9	574	2	US-09-252-991A-30868	Sequence 30868, A
21	61	22.8	170	2	US-09-252-991A-20706	Sequence 20706, A
22	61	22.8	208	2	US-09-252-991A-17849	Sequence 17849, A
23	61	22.8	679	2	US-09-252-991A-27111	Sequence 27111, A
24	60.5	22.6	155	2	US-09-252-991A-32893	Sequence 32893, A
25	60.5	22.6	240	2	US-09-252-991A-28112	Sequence 28112, A
26	60	22.4	343	2	US-09-252-991A-27631	Sequence 27631, A
27	60	22.4	755	2	US-09-902-540-11169	Sequence 11169, A

28	59.5	22.2	158	2	US-09-252-991A-24896	Sequence 24896, A
29	59.5	22.2	293	2	US-09-252-991A-32060	Sequence 32060, A
30	59.5	22.2	395	2	US-09-270-767-44351	Sequence 44351, A
31	59.5	22.2	413	2	US-09-604-231-32	Sequence 32, Appli
32	59.5	22.2	536	2	US-09-359-167-8	Sequence 8, Appli
33	59.5	22.2	606	2	US-09-486-382B-11	Sequence 11, Appli
34	59.5	22.2	683	2	US-09-604-231-30	Sequence 30, Appli
35	59.5	22.2	1031	2	US-09-252-991A-18365	Sequence 18365, A
36	59	22.0	162	2	US-09-252-991A-32101	Sequence 32101, A
37	59	22.0	439	2	US-09-252-991A-17127	Sequence 17127, A
38	59	22.0	770	2	US-09-784-316-5	Sequence 5, Appli
39	59	22.0	770	2	US-10-229-124-5	Sequence 5, Appli
40	58.5	21.8	141	2	US-09-252-991A-23685	Sequence 23685, A
41	58.5	21.8	377	2	US-09-252-991A-32966	Sequence 32966, A
42	58.5	21.8	515	2	US-09-252-991A-23632	Sequence 23632, A
43	58.5	21.8	802	2	US-09-252-991A-25050	Sequence 25050, A
44	58.5	21.8	929	2	US-09-252-991A-19203	Sequence 19203, A
45	58	21.6	117	2	US-09-489-847-199	Sequence 199, App

ALIGNMENTS

RESULT 1

US-09-042-225-2
; Sequence 2, Application US/09042225A
; Patent No. 6207812
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 04930/021001
; CURRENT APPLICATION NUMBER: US/09/042,225A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-042-225-2

Query Match 100.0%; Score 268; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.9e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGPRGAPCRAGAPTIVLTSGRQTLSHGSSSPARATLGKPLVNDYSLN 52
|||||
DB 1 MAAGPRGAPCRAGAPTIVLTSGRQTLSHGSSSPARATLGKPLVNDYSLN 52

RESULT 2

US-09-042-225-8
; Sequence 8, Application US/09042225A
; Patent No. 6207812
; GENERAL INFORMATION:
; APPLICANT: Terek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 04930/021001
; CURRENT APPLICATION NUMBER: US/09/042,225A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-042-225-8

Query Match 27.6%; Score 74; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 RQQTLSHGSSSPARA 38


```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31167

Query Match      24.3%; Score 65; DB 2; Length 205;
Best Local Similarity 32.1%; Pred. No. 1.8;
Matches 18; Conservative 6; Mismatches 18; Indels 14; Gaps 2;

QY 2 AAGPRGAPCRAGATTIV-----LTSGRRQTLSHGSSSP-----ARATLGKP 43
DB 142 APGPGAPFVRGGPPAVRGECFFVRRRAGRRDRASGAAPGCGDRAAGRAGR 197

RESULT 8
US-09-252-991A-25768
; Sequence 25768, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25768
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25768

Query Match      24.3%; Score 65; DB 2; Length 345;
Best Local Similarity 40.5%; Pred. No. 3.4;
Matches 15; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

QY 5 PRPGAPCRAGATTIVLTSGRRQTLSHGSSSPARATLG 41
DB 150 PTPGAPTRIG----LFCRRRRRSTHWRISPSRPTWG 182

RESULT 9
US-09-252-991A-25930
; Sequence 25930, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25930
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25930

Query Match      24.1%; Score 64.5; DB 2; Length 146;
Best Local Similarity 35.4%; Pred. No. 1.4;
Matches 17; Conservative 5; Mismatches 19; Indels 7; Gaps 1;

QY 3 AGRPGAPCRAGATTIVLTSGRRQTLSHGS-----SSPARATLGKP 43
DB 8 ASRPSGSPMTAGRWISIPMGSTRTSTATSSPPAGYSRSPFPSTAVASP 55

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30866
; Sequence 30866, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30866
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30866

Query Match      24.1%; Score 64.5; DB 2; Length 149;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 15; Conservative 2; Mismatches 10; Indels 9; Gaps 1;

QY 4 GPRPGAPCRAGATTIVLTSGRRQTLSHGSSSPARAT 39
DB 99 GARGGSCR-----NGRRQSGHAWGSPQAS 125

RESULT 11
US-09-252-991A-28293
; Sequence 28293, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28293
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28293

Query Match      23.7%; Score 63.5; DB 2; Length 721;
Best Local Similarity 36.8%; Pred. No. 13;
Matches 21; Conservative 4; Mismatches 17; Indels 15; Gaps 2;

QY 4 GRRP-----GAPCRAGATTIVLTSGRRQTLSHG-----SSPARATLGKPLV 45
DB 242 GRRPDPRQAGAPCTARATTGRAGSGRRLTSRPGCRRWRTCGRPANACRRSLGSPSV 298

RESULT 12
US-10-104-047-3112
; Sequence 3112, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
```

```
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3112
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3112

Query Match      23.5%; Score 63; DB 2; Length 128;
Best Local Similarity 35.6%; Pred. No. 1.9;
Matches 16; Conservative 4; Mismatches 23; Indels 2; Gaps 1;

QY 4 GPRGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLKGKPLVND 48
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GPAAGHAAPRGPPVPATAG--AALRAGASEPRQTQGLAPLAGE 44

RESULT 13
US-09-252-991A-19955
; Sequence 19955, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19955
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19955

Query Match      23.5%; Score 63; DB 2; Length 613;
Best Local Similarity 41.0%; Pred. No. 12;
Matches 16; Conservative 5; Mismatches 16; Indels 2; Gaps 1;

QY 5 PRPGAPCRAGAPTIVLTSGRRQTLSHGSSSPARATLKGK 43
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 165 PRPGADPRGAPQAARIAGRRH--AHGFSANPDAPLGRP 201

RESULT 14
US-09-252-991A-22330
; Sequence 22330, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22330
; LENGTH: 1129
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22330

Query Match      23.5%; Score 63; DB 2; Length 1129;
Best Local Similarity 44.8%; Pred. No. 25;
Matches 13; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 13 AGAPTIVLTSGRRQTLSHGSSSPARATLKG 41
      :||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 938 SGSGLVLASGARIDURHGTAAPAEQVDG 966

RESULT 15
US-09-252-991A-25404
; Sequence 25404, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25404
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25404

Query Match      23.1%; Score 62; DB 2; Length 256;
Best Local Similarity 33.3%; Pred. No. 5.7;
Matches 20; Conservative 6; Mismatches 14; Indels 20; Gaps 3;

QY 2 AAGPRPGA-----PC-----RAGAPTIVLTSGRRQTLSHGSSSPARATLKG 42
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 ANGCRPGAALRPVCGAPGRWPCRIAGRGPGAP-VAGTAGRRRTASRARPPPADRGR 142

RESULT 16
US-09-252-991A-19582
; Sequence 19582, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19582
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19582

Query Match      23.1%; Score 62; DB 2; Length 328;
Best Local Similarity 40.0%; Pred. No. 7.7;
Matches 16; Conservative 3; Mismatches 13; Indels 8; Gaps 2;

QY 5 PRPG--APCRAGAPTIVLTSGRRQTLSHGSSSPARATLKG 42
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 PRPGPGPAQAGQP-----ATGQTLRHASLIQPALQAVGR 106

RESULT 17
US-09-252-991A-26099
```

```
; Sequence 26099, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26099
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26099

Query Match      23.1%; Score 62; DB 2; Length 863;
Best Local Similarity 42.4%; Pred. No. 24;
Matches 14; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

QY 5 PRPGAPCRAGAPTIVLTSGRRQTLGSHGSSPAR 37
Db 614 PRGPRWRAGAPAGAGGSRRLRHPAHPQR 646

RESULT 18
US-09-949-016-8300
; Sequence 8300, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8300
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8300

Query Match      22.9%; Score 61.5; DB 2; Length 146;
Best Local Similarity 41.9%; Pred. No. 3.4;
Matches 18; Conservative 3; Mismatches 21; Indels 1; Gaps 1;

QY 2 AAGPRGAPCRAGAPTIVLTSGRRQTLGSHGSSPARATLGKP 43
Db 103 ATGPTSPPPRAAATPATSFGKEDRGLFSIQAPAGLSLP 145

RESULT 19
US-09-252-991A-19910
; Sequence 19910, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
```

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19910
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19910

Query Match      22.9%; Score 61.5; DB 2; Length 367;
Best Local Similarity 42.5%; Pred. No. 10;
Matches 17; Conservative 5; Mismatches 15; Indels 3; Gaps 2;

QY 3 AGPRGAPCRAGA-PTIVLTSGRRQTLGSHGSSPARATLG 41
Db 311 AAPRGRPCAGRALPTPV--PSKPAAYNTGASRPAPASRG 348

RESULT 20
US-09-252-991A-30868
; Sequence 30868, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30868
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30868

Query Match      22.9%; Score 61.5; DB 2; Length 574;
Best Local Similarity 43.2%; Pred. No. 17;
Matches 16; Conservative 2; Mismatches 14; Indels 5; Gaps 1;

QY 5 PRPGAPCRAGAPTIVLTSGRRQTLGSHGSSPARATLG 41
Db 388 PAPWRDCPAGAPGAV-----RDSRPAGTGCPRRATPG 419

RESULT 21
US-09-252-991A-20706
; Sequence 20706, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20706
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20706
```


; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 27631
 ; LENGTH: 343
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27631

Query Match 22.4%; Score 60; DB 2; Length 343;
 Best Local Similarity 41.3%; Pred. No. 14;
 Matches 19; Conservative 5; Mismatches 12; Indels 10; Gaps 4;

QY 5 PRPGAPCRAGAP-TIVLTSGRR--QTL-----SHGSSSPARATLQK 43
 DB 169 PHGCRP---GLPREVLQSGRRHLQDLADPPHTGSGSVSAGTSGSP 211

RESULT 27
 US-09-540-11169
 ; Sequence 11169, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 11169
 ; LENGTH: 755
 ; TYPE: PRT
 ; ORGANISM: Myxococcus xanthus
 US-09-540-11169

Query Match 22.4%; Score 60; DB 2; Length 755;
 Best Local Similarity 30.4%; Pred. No. 37;
 Matches 14; Conservative 7; Mismatches 19; Indels 6; Gaps 1;

QY 1 MAAGPRPGAPCRAGAP-TIVLTSGRRQTLSHGSSSPARATLQKPLVL 46
 DB 365 LPALPGGAPPGVPPPLDAAASRTWGM-----PRAVLGAPIIV 404

RESULT 28
 US-09-252-991A-24896
 ; Sequence 24896, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 24896
 ; LENGTH: 158
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-24896

Query Match 22.2%; Score 59.5; DB 2; Length 158;

Best Local Similarity 32.7%; Pred. No. 6.6;
 Matches 16; Conservative 5; Mismatches 21; Indels 7; Gaps 1;
 QY 2 AAGPRPGAPCRAGA-----PTIVLTSGRRQTLSHGSSSPARATLQK 43
 DB 28 SAGAMPGAPRMAGSRSSGRPSNATPASSRNRLRQPNQNSIRVPSGTP 76

RESULT 29
 US-09-252-991A-32060
 ; Sequence 32060, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 32060
 ; LENGTH: 293
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-32060

Query Match 22.2%; Score 59.5; DB 2; Length 293;
 Best Local Similarity 37.8%; Pred. No. 14;
 Matches 17; Conservative 4; Mismatches 19; Indels 5; Gaps 2;

QY 3 AGPRPGAPCRAGAP-TIVLTSGRRQTLSHGSSSPA-----RATLQK 43
 DB 130 AGPRDGTSPAGSQPVLRLAA-REPDPGQPPAAQRTORAPAGK 173

RESULT 30
 US-09-270-767-44351
 ; Sequence 44351, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 44351
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; FEATURE:
 ; OTHER INFORMATION: Xaa means any amino acid
 US-09-270-767-44351

Query Match 22.2%; Score 59.5; DB 2; Length 395;
 Best Local Similarity 34.0%; Pred. No. 20;
 Matches 18; Conservative 5; Mismatches 19; Indels 11; Gaps 2;

QY 2 AAGPRPGAPCRAGA-----PTIVLTSGRRQTLSHGSSSPARA-----TLQK 43
 DB 163 AGGGAPGLPAKAGATGIPQPAVAAAPGAPATQSAVTTPAAPAPATPQTTP 215

RESULT 31
 US-09-604-231-32
 ; Sequence 32, Application US/09604231
 ; Patent No. 6884614
 ; GENERAL INFORMATION:

```
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: PHOSPHOENOLPYRUVATE:SUGAR PHOSPHOTRANSFERASE
; TITLE OF INVENTION: SYSTEM PROTEINS
; FILE REFERENCE: BGI-122CP
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 32
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-604-231-32
```

```
Query Match 22.2%; Score 59.5; DB 2; Length 413;
Best Local Similarity 39.5%; Pred. No. 21;
Matches 15; Conservative 5; Mismatches 15; Indels 3; Gaps 1;
```

```
Qy 13 AGAPTIVLTSGRRQTLHGSSSPARATLGKPLVNDYS 50
| | | | | : | | | | | : | | | | | : | | | | |
Db 196 AAIPAALLTP---EFLALGSAGDTVTVFGLPVNDYS 230
```

```
RESULT 32
US-09-359-167-8
; Sequence 8, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-359-167-8
```

```
Query Match 22.2%; Score 59.5; DB 2; Length 536;
Best Local Similarity 45.0%; Pred. No. 29;
Matches 18; Conservative 3; Mismatches 12; Indels 7; Gaps 2;
```

```
Qy 1 MAAG---PRPGAPCRAGAPTIVLTSGRRQTLHGSSSPA 36
| | | | | : | | | | | : | | | | | : | | | | |
Db 1 MAAGAMTFRPVQPARPGFGL---SGRRSLLCQVASTPA 37
```

```
RESULT 33
US-09-486-382B-11
; Sequence 11, Application US/09486382B
; Patent No. 6388174
; GENERAL INFORMATION:
; APPLICANT: Hokko Chemical Industry Co., Ltd.
; APPLICANT: MINISTRY OF AGRICULTURE, FORESTRY AND FISHERIES
; TITLE OF INVENTION: Genes for encoding alpha-subunits of anthranilate
; TITLE OF INVENTION: synthase of rice, and DNAs related to said genes
; FILE REFERENCE: 19647
; CURRENT APPLICATION NUMBER: US/09/486,382B
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: JP 9-235049
; PRIOR FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 11
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-486-382B-11
```

```
Query Match 22.2%; Score 59.5; DB 2; Length 606;
Best Local Similarity 38.1%; Pred. No. 33;
Matches 16; Conservative 2; Mismatches 17; Indels 7; Gaps 2;
```

```
Qy 6 RCGAPCRAGAP---TIVLTSGRRQTLHGSSSPARATLGKP 43
| | | | | : | | | | | : | | | | | : | | | | |
Db 17 RPATPAAAAAPVRAAAVAGRRRTTSRRGG---VRCSAGKP 55
```

```
RESULT 34
US-09-604-231-30
; Sequence 30, Application US/09604231
; Patent No. 6884614
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: PHOSPHOENOLPYRUVATE:SUGAR PHOSPHOTRANSFERASE
; FILE REFERENCE: BGI-122CP
; CURRENT APPLICATION NUMBER: US/09/604,231
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 30
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-604-231-30
```

```
Query Match 22.2%; Score 59.5; DB 2; Length 683;
Best Local Similarity 39.5%; Pred. No. 38;
Matches 15; Conservative 5; Mismatches 15; Indels 3; Gaps 1;
```

```
Qy 13 AGAPTIVLTSGRRQTLHGSSSPARATLGKPLVNDYS 50
| | | | | : | | | | | : | | | | | : | | | | |
Db 196 AAIPAALLTP---EFLALGSAGDTVTVFGLPVNDYS 230
```

```
RESULT 35
US-09-252-991A-18365
; Sequence 18365, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18365
; LENGTH: 1031
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18365
```

```
Query Match 22.2%; Score 59.5; DB 2; Length 1031;
Best Local Similarity 33.3%; Pred. No. 63;
Matches 17; Conservative 3; Mismatches 20; Indels 11; Gaps 1;
```

```
QY 4 GPRGAPCRGAP-----TIVLTSGRRQTLSHGSSSPARATLQKP 43
Db 621 GPDGNPCGAPGHRRTGPRGRLRRAGGRLPALPSGAPQPSRADRVRP 671

RESULT 36
US-09-252-991A-32101
; Sequence 32101, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32101
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32101

Query Match 22.0%; Score 59; DB 2; Length 162;
Best Local Similarity 39.0%; Pred. No. 7.9;
Matches 16; Conservative 4; Mismatches 17; Indels 4; Gaps 2;

QY 5 PR--PGAPCRGAPRTIVLTSGRRQTLSHGSSSPARATLQKP 43
Db 60 PRTTTHAPCRGASAPPAHRAHRSVAGPGRS--ARPVGRP 98

RESULT 37
US-09-252-991A-17127
; Sequence 17127, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17127
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17127

Query Match 22.0%; Score 59; DB 2; Length 439;
Best Local Similarity 48.3%; Pred. No. 26;
Matches 14; Conservative 3; Mismatches 8; Indels 4; Gaps 1;

QY 1 MAAGPRPGAPCR-----AGAPTIVLTSGRR 25
Db 146 LLAGRPGVPRRAVAPAGPVLLRRGSR 174

RESULT 38
US-09-784-316-5
; Sequence 5, Application US/09784316
; Patent No. 6461843
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
```

```
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001139
; CURRENT APPLICATION NUMBER: US/09/784,316
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-784-316-5

Query Match 22.0%; Score 59; DB 2; Length 770;
Best Local Similarity 37.1%; Pred. No. 51;
Matches 26; Conservative 3; Mismatches 21; Indels 20; Gaps 5;

QY 3 AGP-----RPGAPCRGAP-----TIVLTSGRRQTLSHGSSSPARAT-----LQK 42
Db 523 AGPQNKHKRPGEPSPMQGIPKRRRCASPLLRGRQSPAVNSHIGGKPPAPMTQAQPGLIK 582

QY 43 PLVLDYSLN 52
Db 583 PLPLHKEATN 592

RESULT 39
US-10-229-124-5
; Sequence 5, Application US/10229124
; Patent No. 6723542
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001139DIV
; CURRENT APPLICATION NUMBER: US/10/229,124
; CURRENT FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-229-124-5

Query Match 22.0%; Score 59; DB 2; Length 770;
Best Local Similarity 37.1%; Pred. No. 51;
Matches 26; Conservative 3; Mismatches 21; Indels 20; Gaps 5;

QY 3 AGP-----RPGAPCRGAP-----TIVLTSGRRQTLSHGSSSPARAT-----LQK 42
Db 523 AGPQNKHKRPGEPSPMQGIPKRRRCASPLLRGRQSPAVNSHIGGKPPAPMTQAQPGLIK 582

QY 43 PLVLDYSLN 52
Db 583 PLPLHKEATN 592

RESULT 40
US-09-252-991A-23685
; Sequence 23685, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
```


Thu Dec 29 19:53:29 2005

```

RESULT 45
US-09-489-847-199
; Sequence 199, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-199

```

```

Query Match      21.6%; Score 58; DB 2; Length 117;
Best Local Similarity 30.4%; Pred.No. 7.2;
Matches 17; Conservative 4; Mismatches 19; Indels 16; Gaps 1;

QY      1 MAAGPRPGA-----PCRAGAPTIVLTGRROTLSHGSSSPARATL 40
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      18 LAAGTPSTGSSSPSWKHIGTSLOKTRGSLFTTILTSAGAGQSTGTGNPAAGRSL 73

```

Search completed: December 27, 2005, 23:05:52
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 23:04:24 ; Search time 163 Seconds
(without alignments)
133.295 Million cell updates/sec

Title: US-09-819-144A-2

Perfect score: 268

Sequence: 1 MAAGPRPGAPCRAGAPTIVL.....SSPARATLGKPLVNDYSLN 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*

2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*

3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*

4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*

5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*

6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	268	100.0	52	3	US-09-819-144A-2
2	79	29.5	229	4	US-10-425-115-222906
3	77	28.7	190	4	US-10-437-963-198300
4	74	27.6	16	3	US-09-819-144A-8
5	73.5	27.4	791	4	US-10-170-385-57
6	73.5	27.4	863	4	US-10-359-012-2
7	73.5	27.4	863	4	US-10-359-012-14
8	71.5	26.7	240	4	US-10-425-115-226148
9	71	26.5	19608	4	US-10-084-846A-8
10	68	25.4	315	4	US-10-080-170-423
11	68	25.4	315	4	US-10-080-170-423
12	68	25.4	315	4	US-10-468-356-423
13	67	25.0	219	4	US-10-425-115-262578
14	67	25.0	1321	5	US-10-840-512-117
15	66.5	24.8	185	4	US-10-425-115-341679
16	65.5	24.4	923	5	US-10-756-149-5246
17	64.5	24.1	137	4	US-10-767-701-33507
18	64.5	24.1	309	5	US-10-450-763-43334
19	64.5	24.1	596	4	US-10-156-761-8931
20	64	23.9	185	4	US-10-425-115-270067
21	64	23.9	359	4	US-10-437-963-190864
22	63.5	23.7	74	4	US-10-425-115-263976
23	63.5	23.7	87	4	US-10-425-115-313357
24	63.5	23.7	223	4	US-10-425-114-62626
25	63	23.5	128	4	US-10-104-047-3112
26	63	23.5	128	4	US-10-108-260A-2594
27	62.5	23.3	118	4	US-10-424-599-271704

28 62.5 23.3 197 4 US-10-437-963-128246 Sequence 128246,
29 62.5 23.3 208 4 US-10-437-963-176133 Sequence 176133,
30 62.5 23.3 524 5 US-10-450-763-58710 Sequence 38710, A
31 62 23.1 79 4 US-10-425-115-259836 Sequence 259836,
32 62 23.1 175 4 US-10-425-115-210919 Sequence 210919,
33 62 23.1 177 4 US-10-437-963-109208 Sequence 109208,
34 62 23.1 259 4 US-10-443-622-59 Sequence 59, Appl
35 62 23.1 259 5 US-10-974-440-30 Sequence 30, Appl
36 61.5 22.9 135 4 US-10-425-115-350149 Sequence 350149,
37 61.5 22.9 154 4 US-10-767-701-55508 Sequence 55508, A
38 61.5 22.9 170 3 US-09-864-408A-6684 Sequence 6684, Ap
39 61.5 22.9 240 4 US-10-425-115-326091 Sequence 326091,
40 61 22.8 79 4 US-10-425-115-357643 Sequence 357643,
41 61 22.8 169 4 US-10-767-701-42973 Sequence 42973, A
42 61 22.8 176 4 US-10-425-115-203617 Sequence 203617,
43 61 22.8 241 4 US-10-425-115-245202 Sequence 245202,
44 61 22.8 595 4 US-10-156-761-10072 Sequence 10072, A
45 61 22.8 656 4 US-10-108-260A-2611 Sequence 2611, Ap

ALIGNMENTS

RESULT 1

US-09-819-144A-2

; Sequence 2, Application US/09819144A

; Publication No. US20010016649A1

; GENERAL INFORMATION:

; APPLICANT: Terek, Richard M.

; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES

; FILE REFERENCE: 21486-021DIV

; CURRENT APPLICATION NUMBER: US/09/819,144A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 52

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-819-144A-2

Query Match 100.0%; Score 268; DB 3; Length 52;

Best Local Similarity 100.0%; Pred. No. 2.6e-24;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLTSHGSSSPARATLGKPLVNDYSLN 52

DB 1 MAAGPRPGAPCRAGAPTIVLTSGRRQTLTSHGSSSPARATLGKPLVNDYSLN 52

RESULT 2

US-10-425-115-222906

; Sequence 222906, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE OF INVENTION: Plance

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 222906

; LENGTH: 229

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(229)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_134881C.1.pep
US-10-425-115-222906

Query Match 29.5%; Score 79; DB 4; Length 229;
Best Local Similarity 43.8%; Pred. No. 0.38; Indels 6; Gaps 2;
Matches 21; Conservative 5; Mismatches 16;

Qy 4 GPRPGACRAG--AFTVILTSGRRQTLSHG----SSSPARATLKGKPLV 45
Db 32 GOLPAKVCRTGSAPTRLXRRRRRTLHGGSPELSSPPPTICRLL 79

RESULT 3

US-10-437-963-198300
; Sequence 198300, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198300
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93974C.1.pep
US-10-437-963-198300

Query Match 28.7%; Score 77; DB 4; Length 190;
Best Local Similarity 35.1%; Pred. No. 0.54; Indels 10; Gaps 3;
Matches 20; Conservative 11; Mismatches 16;

Qy 1 MAAGPRGAC-----CRAGAPTIVLTSGRRQTLSHGSSSPARATLKGKPLVINDYSL 51
Db 131 LAPAP-PCGAPSPGITICSGGPTVMTPTG---VLSYGAASRSSANLSSLLIVANVSL 183

RESULT 4

US-09-819-144A-8
; Sequence 8, Application US/09819144A
; Publication No. US20010016649A1
; GENERAL INFORMATION:
; APPLICANT: Tersek, Richard M.
; TITLE OF INVENTION: CHONDROSARCOMA ASSOCIATED GENES
; FILE REFERENCE: 21486-021DIV
; CURRENT APPLICATION NUMBER: US/09/819,144A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-144A-8

Query Match 27.6%; Score 74; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.078; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;

Qy 24 RRQTLSHGSSSPARA 38
; |||||

Db 1 RRQTLSHGSSSPARA 15

RESULT 5

US-10-170-385-57
; Sequence 57, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-57

Query Match 27.4%; Score 73.5; DB 4; Length 791;
Best Local Similarity 45.7%; Pred. No. 6.8; Indels 1; Gaps 1;
Matches 16; Conservative 5; Mismatches 13;

Qy 2 AAGPRGACRAGAPTIVLTSGRRQTL-SHGSSSP 35
Db 516 SSGPQPAKPCSGATPTLLLVGDRSPVPSPGSSSP 550

RESULT 6

US-10-359-012-2
; Sequence 2, Application US/10359012
; Publication No. US20030232419A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: KOLODKIN, Alex L.
; APPLICANT: TERMAN, Jon R.
; APPLICANT: MAO, Tianyi
; APPLICANT: PASTERKAMP, Ronald J.
; APPLICANT: YU, Hung-Hsiang
; TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
; FILE REFERENCE: JHU1840-3
; CURRENT APPLICATION NUMBER: US/10/359,012
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/388,325
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/384,302
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/354,178
; PRIOR FILING DATE: 2002-02-04
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-359-012-2


```
Query Match      27.4%; Score 73.5; DB 4; Length 863;
Best Local Similarity 45.7%; Pred. No. 7.5;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY  2 AAGPRPGACAGATTIVLTSGRQTL-SHGSSSP 35
Db   SSGPQAPKPCGATPTLLVGDSPVPSPGSSSP 622

RESULT 7
US-10-359-012-14
; Sequence 14, Application US/10359012
; Publication No. US20030232419A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: KOLODKIN, Alex L.
; APPLICANT: TERMAN, Jon R.
; APPLICANT: MAO, Tianyi
; APPLICANT: PASTERKAMP, Ronald J.
; APPLICANT: YU, Hung-Hsiang
; TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
; TITLE OF INVENTION: AND METHODS OF USING THE SAME
; FILE REFERENCE: JH1840-3
; CURRENT APPLICATION NUMBER: US/10/359,012
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/388,325
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/384,302
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/354,178
; PRIOR FILING DATE: 2002-02-04
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-359-012-14

Query Match      27.4%; Score 73.5; DB 4; Length 863;
Best Local Similarity 45.7%; Pred. No. 7.5;
Matches 16; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY  2 AAGPRPGACAGATTIVLTSGRQTL-SHGSSSP 35
Db   SSGPQAPKPCGATPTLLVGDSPVPSPGSSSP 622

RESULT 8
US-10-425-115-226148
; Sequence 226148, Application US/10425115
; Publication No. US2004021427A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 226148
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_137836C.1.pap
US-10-425-115-226148

Query Match      26.7%; Score 71.5; DB 4; Length 240;
Best Local Similarity 35.1%; Pred. No. 3.1;

Matches 26; Conservative 4; Mismatches 19; Indels 25; Gaps 3;

QY  1 MAAGPRPGA-PCRAGA-----PTIVLTSGRRQTL-SHGSSSPA 36
Db   LAGPRPGAGPPTTFAADPRVDSCTVAHFSCPAPSPHPSPQLLS-RRRLPSGGPRSPR 122

QY  37 RATLGKPLVLDYS 50
Db   TATGKPLAAHDIS 136

RESULT 9
US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREFZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 8
; LENGTH: 19608
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-8

Query Match      26.5%; Score 71; DB 4; Length 19608;
Best Local Similarity 39.6%; Pred. No. 4.8e+02;
Matches 19; Conservative 2; Mismatches 15; Indels 12; Gaps 2;

QY  2 AAGPRPGACAGATTIVLTSGRQTL-SHGSSSPARATLGKP 43
Db   ARPPRPGHCRGAP-----GKSGRGRGTAALADSSPGRTARSGP 3682

RESULT 10
US-10-080-170-423
; Sequence 423, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 423
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-423

Query Match      25.4%; Score 68; DB 4; Length 315;
Best Local Similarity 28.2%; Pred. No. 11;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;
```

QY 1 MAAGPRGAPC-----RAGAPT-----IVLTSGRRQTL 28
: ||| ||| : : : : :
Db 51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAEVLVHPQSRRRGIGTAMARAALAKTAGRNQFW 110
: ||| ||| : : : : :
QY 29 SHGSSSPARAT 39
: ||| ||| : : : : :
Db 111 AHGTLDPARAT 121

RESULT 11
US-10-080-170-423
; Sequence 423, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 423
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-423

Query Match 25.4%; Score 68; DB 4; Length 315;
Best Local Similarity 28.2%; Pred. No. 11;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;

QY 1 MAAGPRGAPC-----RAGAPT-----IVLTSGRRQTL 28
: ||| ||| : : : : :
Db 51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAEVLVHPQSRRRGIGTAMARAALAKTAGRNQFW 110
: ||| ||| : : : : :
QY 29 SHGSSSPARAT 39
: ||| ||| : : : : :
Db 111 AHGTLDPARAT 121

RESULT 12
US-10-468-356-423
; Sequence 423, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn ver. 3.2
; SEQ ID NO 423
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-423

Query Match 25.4%; Score 68; DB 4; Length 315;
Best Local Similarity 28.2%; Pred. No. 11;
Matches 20; Conservative 5; Mismatches 14; Indels 32; Gaps 2;

QY 1 MAAGPRGAPC-----RAGAPT-----IVLTSGRRQTL 28
: ||| ||| : : : : :
Db 51 LVAGSRPGGPIIGYLNLSPPRGAGGAMAEVLVHPQSRRRGIGTAMARAALAKTAGRNQFW 110
: ||| ||| : : : : :
QY 29 SHGSSSPARAT 39
: ||| ||| : : : : :
Db 111 AHGTLDPARAT 121

RESULT 13
US-10-425-115-262578
; Sequence 262578, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 262578
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_171081C.1.pep
US-10-425-115-262578

Query Match 25.0%; Score 67; DB 4; Length 219;
Best Local Similarity 38.1%; Pred. No. 9.7;
Matches 24; Conservative 3; Mismatches 20; Indels 16; Gaps 4;

QY 3 AGPRPGAP-CRAGAPTIVLTSGRRQTLSHGS-----SSPA-----RATLGKPLVLN 47
: ||| ||| : : : : :
Db 152 AGLAPGAPSCRKGCRAARTGGPPKT-SHAARAFAFVWGPGFAPFVRRLRAEVGAPLVRE 210
: ||| ||| : : : : :
QY 48 DYS 50
: ||| : : : : :
Db 211 DLS 213

RESULT 14
US-10-840-512-117
; Sequence 117, Application US/10840512
; Publication No. US20050125852A1
; GENERAL INFORMATION:
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1455
; CURRENT APPLICATION NUMBER: US/10/840,512
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469,014
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 117
; LENGTH: 1321
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-840-512-117

Query Match 25.0%; Score 67; DB 5; Length 1321;
Best Local Similarity 39.5%; Pred. No. 72;
Matches 17; Conservative 8; Mismatches 10; Indels 8; Gaps 2;

QY 7 PGAPCRAGAPTIVL-----TSGRRQTLSHGSSSPARATLGKPL 44


```
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8931
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8931

Query Match      24.1%; Score 64.5; DB 4; Length 596;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 18; Conservative 3; Mismatches 10; Indels 5; Gaps 1;

Qy 11 CR-----AGATTIVLTGRRQTLGSHSSSPARATIG 41
Db 288 CRREDSAGAVFRVRLIGRLQFTSAGSASPARSLG 323

RESULT 20
US-10-425-115-270067
; Sequence 270067, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 270067
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_177899C.1.pep
US-10-425-115-270067

Query Match      23.9%; Score 64; DB 4; Length 165;
Best Local Similarity 38.2%; Pred. No. 16;
Matches 13; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

Qy 4 GPRGACRAGAPTIVLTGRRQTLGSHSSSPAR 37
Db 83 GPSRCFCRCRGPRPRSSRGPRTPASGTEASRR 116

RESULT 21
US-10-437-963-190864
; Sequence 190864, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190864
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_87238C.1.pep
US-10-437-963-190864

Query Match      23.9%; Score 64; DB 4; Length 359;
Best Local Similarity 30.0%; Pred. No. 38;
Matches 24; Conservative 6; Mismatches 14; Indels 36; Gaps 4;

Qy 2 AAGPRPGACRAGA-PTIVLTG-----GRRQ-----T 27
Db 176 AAAPRSGSTGRRGARPSIALTNELGKRRERAPHLHPQRP RRGGTYTTTAAATFQP VQ 235
Qy 28 LSHGSSSPARATIGKPLV LN 47
Db 236 LSHSPNPRTPL--PFLD 253

RESULT 22
US-10-425-115-263976
; Sequence 263976, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 263976
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_172359C.1.pep
US-10-425-115-263976

Query Match      23.7%; Score 63.5; DB 4; Length 74;
Best Local Similarity 48.6%; Pred. No. 7.6;
Matches 17; Conservative 0; Mismatches 9; Indels 9; Gaps 2;

Qy 3 AGPR--PGACRAGAPTIVLTGRRQTLGSHSSSP 35
Db 2 AGPRAPGVVRAGA-----GRRAGPRHGPREP 29

RESULT 23
US-10-425-115-313357
; Sequence 313357, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
```

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 313357
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_4883C.1.pep
US-10-425-115-313357

Query Match      23.7%; Score 63.5; DB 4; Length 87;
Best Local Similarity 33.3%; Pred. No. 9.1;
Matches 17; Conservative 4; Mismatches 13; Indels 17; Gaps 3;

QY 5 PRPGAP-----CRAGPTVLTSGRRQTLSHGSSSPARATLGKP 43
Db 21 PNFAGPNYSKTPQTRACKRGPPPL-----GRGE-LGDGLGAPGNGVLKP 66

RESULT 24
US-10-425-114-62626
; Sequence 62626, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62626
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-112-D11_FLI.pep
US-10-425-114-62626

Query Match      23.7%; Score 63.5; DB 4; Length 223;
Best Local Similarity 34.8%; Pred. No. 26;
Matches 16; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

QY 2 AAGPRGACR-AGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVL 46
Db 106 SAGPRPSSTSRPPWAAAAAARRRLLPRGAASPRRRARPLL 151

RESULT 25
US-10-104-047-3112
; Sequence 3112, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3112
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3112
```

```
Query Match      23.5%; Score 63; DB 4; Length 128;
Best Local Similarity 35.6%; Pred. No. 16;
Matches 16; Conservative 4; Mismatches 23; Indels 2; Gaps 1;

QY 4 GPRPGACRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLND 48
Db 2 GPAAGHAAPRGPPVPATTAG--AALRAGASEPRQTQGLGAPLALGE 44

RESULT 26
US-10-108-260A-2594
; Sequence 2594, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2594
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2594

Query Match      23.5%; Score 63; DB 4; Length 128;
Best Local Similarity 35.6%; Pred. No. 16;
Matches 16; Conservative 4; Mismatches 23; Indels 2; Gaps 1;

QY 4 GPRPGACRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLVLND 48
Db 2 GPAAGHAAPRGPPVPATTAG--AALRAGASEPRQTQGLGAPLALGE 44

RESULT 27
US-10-424-599-271704
; Sequence 271704, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 271704
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87368C.1.pep
US-10-424-599-271704

Query Match      23.3%; Score 62.5; DB 4; Length 118;
Best Local Similarity 38.3%; Pred. No. 17;
Matches 18; Conservative 5; Mismatches 19; Indels 5; Gaps 1;

QY 2 AAGPRPGACRAGAPTIVLTSGRRQTL-----HGSSSPARATLGKP 43
Db 15 APRPPGAPPPGSGPPPGGLGSPRSGACAPPGCGGASAPARASVGAP 61

RESULT 28
US-10-437-963-128246
; Sequence 128246, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 128246
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(197)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30618C.1.pep
US-10-437-963-128246

Query Match 23.3%; Score 62.5; DB 4; Length 197;
Best Local Similarity 39.5%; Pred. No. 30;
Matches 17; Conservative 4; Mismatches 17; Indels 5; Gaps 2;

Qy 2 AAGPRGAPCGAGTIVLTSGRRQTLHGSSSPAR--ATLQK 42
Db 3 AAGPSSAPPRPGAGVATMAAGR---GHGDGDGAALGAAYGR 42

RESULT 29

US-10-437-963-176133
; Sequence 176133, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 176133
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73911C.1.pep
US-10-437-963-176133

Query Match 23.3%; Score 62.5; DB 4; Length 208;
Best Local Similarity 45.0%; Pred. No. 31;
Matches 18; Conservative 1; Mismatches 16; Indels 5; Gaps 2;

Qy 5 PRGAPCEA-CAPTIVLTSGRRQTLHGSSSPARATLQKP 43
Db 58 PPQAPCRGRRLPTTL----RSAPSSSSMPPPRATLDLP 93

RESULT 30

US-10-450-763-38710

; Sequence 38710, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 38710
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (120)..(458)
; OTHER INFORMATION: AIR synthase related protein domain identified by Pfam.
; OTHER INFORMATION: accession name AIRS, E-value=1.8e-06, Pfam score of -5.1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(524)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-38710

Query Match 23.3%; Score 62.5; DB 5; Length 524;
Best Local Similarity 44.7%; Pred. No. 88;
Matches 17; Conservative 3; Mismatches 11; Indels 7; Gaps 2;

Qy 10 PCRAGTIVLTSGRRQTLHGSSSPARATLQK 42
Db 9 PCRAGA--MAETSATGACGEAMAAEGSGPAGTLGR 44

RESULT 31

US-10-425-115-259836
; Sequence 259836, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 259836
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_168589C.1.pep
US-10-425-115-259836

Query Match 23.1%; Score 62; DB 4; Length 79;
Best Local Similarity 37.8%; Pred. No. 12;
Matches 14; Conservative 7; Mismatches 14; Indels 2; Gaps 1;

Qy 7 PGAPCRAGTIVLTSGRRQTLHGSSSPARATLQKP 43
Db 6 PSSPCRVSAVSTAQTPRLFEVFS--ARAPARRTCGKP 40

RESULT 32

```
US-10-425-115-210919
; Sequence 210919, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 210919
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_123954C.1.pap
US-10-425-115-210919

Query Match      23.1%; Score 62; DB 4; Length 175;
Best Local Similarity 45.2%; Pred. No. 30;
Matches 14; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY      10 PCRAGAPCTIVLTSGRRQTLTSHGSSSPARATL 40
Db      78 PTSAAPPSLWSSPGRLQPHAGHSSPPHRAPL 108

RESULT 33
US-10-437-963-109208
; Sequence 109208, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109208
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(177)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13388C.1.pap
US-10-437-963-109208

Query Match      23.1%; Score 62; DB 4; Length 177;
Best Local Similarity 42.9%; Pred. No. 30;
Matches 15; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

QY      5 PRPGAPCRAGAPCTIVLTSGRRQTLTSHGSSSPAR 37
Db      96 PRPPASARSPPPLAAAGPRATARRHSRSPAR 130

RESULT 34
US-10-443-622-59
```

```
; Sequence 59, Application US/10443622
; Publication No. US20040024192A1
; GENERAL INFORMATION:
; APPLICANT: Carter et al.
; TITLE OF INVENTION: 19 Human Secreted Proteins
; FILE REFERENCE: P2009P1
; CURRENT APPLICATION NUMBER: US/10/443,622
; CURRENT FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: PCT/US98/13608
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 60/051,480
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 60/051,381
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 60/058,663
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,598
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (78)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-443-622-59

Query Match      23.1%; Score 62; DB 4; Length 259;
Best Local Similarity 30.4%; Pred. No. 46;
Matches 21; Conservative 5; Mismatches 19; Indels 24; Gaps 3;

QY      5 PRPGAPCRAGAPCTIVLTSGRR-----QTL-----SHGSSSPARATL----- 40
Db      135 PAPKEPAGPGAPLLVVGSSRNLLPLSVMSASQALQTVALSAHGSSEPNLALKALAFN 194
QY      41 GKPLVLNDY 49
Db      195 GSPLRFDKY 203

RESULT 35
US-10-974-440-30
; Sequence 30, Application US/10974440
; Publication No. US20050214795A1
; GENERAL INFORMATION:
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: REDDY, Roopa
; APPLICANT: TANG, Y. Tom
; APPLICANT: GERSTIN, Edward H.
; APPLICANT: ARVIZU, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: LU, Dyung Aina M.
; TITLE OF INVENTION: Human Transcriptional Regulator Molecules
; FILE REFERENCE: PF-0509 USN
; CURRENT APPLICATION NUMBER: US/10/974,440
; CURRENT FILING DATE: 2004-10-26
; PRIOR APPLICATION NUMBER: US/09/674,743
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/09935
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/084,254
; PRIOR FILING DATE: 1998-05-04
; PRIOR APPLICATION NUMBER: 60/095,827
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: 60/102,745
; PRIOR FILING DATE: 1998-10-02
```

```
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1910984CD1
US-10-974-440-30

Query Match      23.1%; Score 62; DB 5; Length 259;
Best Local Similarity 30.4%; Pred. No. 46;
Matches 21; Conservative 5; Mismatches 19; Indels 24; Gaps 3;

Qy 5 PRGPACRAGAPTIVLTSGRR-----QTL-----SHGSSSPARATL----- 40
Db 135 PAPKEPAGPGLLVGSSRNLLSPLSVMSASQALQIVALSAAHGSSSEPNLAKALAFN 194
Qy 41 GKPLVINDY 49
Db 195 GSPLRFDKY 203

RESULT 36
US-10-425-115-350149
; Sequence 350149, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 350149
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_82508C.1.pep
US-10-425-115-350149

Query Match      22.9%; Score 61.5; DB 4; Length 135;
Best Local Similarity 40.9%; Pred. No. 26;
Matches 18; Conservative 1; Mismatches 22; Indels 3; Gaps 1;

Qy 3 AGPRPGAPCRAGAPTIVLTSGRRQT---LSHGSSSPARATLGKP 43
Db 1 AWPSPGAPPGGGTSSGSAPORTCWGTSARSRPPPRITLGT 44

RESULT 37
US-10-767-701-55508
; Sequence 55508, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 55508
; LENGTH: 154
; TYPE: PRT
```

```
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30163676.pep
US-10-767-701-55508

Query Match      22.9%; Score 61.5; DB 4; Length 154;
Best Local Similarity 38.5%; Pred. No. 30;
Matches 20; Conservative 4; Mismatches 17; Indels 11; Gaps 3;

Qy 3 AGPRPGAPCR---GAPTIVLTSGRR-----QTLSHGSSSPARATLGKPLVL 46
Db 65 AAARPSLPARSRDGSPR---ASGRMKSRSTRSGAPMPPTLPLPLPL 113

RESULT 38
US-09-864-408A-6684
; Sequence 6684, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 6684
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (138)..(138)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-09-864-408A-6684

Query Match      22.9%; Score 61.5; DB 3; Length 170;
Best Local Similarity 32.2%; Pred. No. 33;
Matches 19; Conservative 4; Mismatches 13; Indels 23; Gaps 3;

Qy 6 RPGAP-----CRAG-----APTIVLTSGRR---QTLSHGSSSPARATLG 41
Db 92 RGGIPISLTMTTCRCGRFFPTTLPPLPAASPTVSRPSGRNTSPTSXDGSSAPLRPVG 150

RESULT 39
US-10-425-115-326091
; Sequence 326091, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 326091
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(240)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_60462C.1.pep
```


US-10-425-115-326091

Query Match 22.9%; Score 61.5; DB 4; Length 240;
Best Local Similarity 40.0%; Pred. No. 48;
Matches 20; Conservative 5; Mismatches 14; Indels 11; Gaps 3;

Qy 1 MAAGPRGAPCRAGAPTIVLTSGR-----RQLSHGSSSPARATLGKP 43
Db 69 VAAGRAPPSP---GSPASVLCGRNRIRVPREDLHASLGA-PSLGSP 114

RESULT 40

US-10-425-115-357643
; Sequence 357643, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 357643
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(79)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_8933C.1.pep
US-10-425-115-357643

Query Match 22.8%; Score 61; DB 4; Length 79;
Best Local Similarity 35.7%; Pred. No. 16;
Matches 15; Conservative 6; Mismatches 21; Indels 0; Gaps 0;

Qy 2 AAGPRGAPCRAGAPTIVLTSGRRQLSHGSSSPARATLGKP 43
Db 16 AAGPRKRTGRELHQRVAHGRROEVTTQATGPRVAYQGRP 57

RESULT 41

US-10-767-701-42973
; Sequence 42973, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42973
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(169)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C35069_1.pep
US-10-767-701-42973

Query Match 22.8%; Score 61; DB 4; Length 169;
Best Local Similarity 38.2%; Pred. No. 38;
Matches 21; Conservative 2; Mismatches 20; Indels 12; Gaps 2;

Qy 6 RPAGAP----CRAGAPTIVLTSGRR-----QTLSHGSSSPARATLGKPLVLD 48
Db 30 RPGTPRAPRLRAGATTGPPWRRAFWRRMEAPRLGRXSALPARTAEGLLVFED 84

RESULT 42

US-10-425-115-203617
; Sequence 203617, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 203617
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_117288C.1.pep
US-10-425-115-203617

Query Match 22.8%; Score 61; DB 4; Length 176;
Best Local Similarity 34.5%; Pred. No. 39;
Matches 19; Conservative 5; Mismatches 17; Indels 14; Gaps 2;

Qy 3 AGRPGAPCRAGAPTIVLTSGRRQ-----TLSHGSSSPA-RATLGKP 43
Db 19 AHRPAPPARLLPSSVLSGRRAPPLQCPAPRHSVLTASHPARRAORSALDRP 73

RESULT 43

US-10-425-115-245202
; Sequence 245202, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 245202
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(241)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_155211C.1.pep
US-10-425-115-245202

Query Match 22.8%; Score 61; DB 4; Length 241;
Best Local Similarity 39.5%; Pred. No. 56;
Matches 17; Conservative 3; Mismatches 13; Indels 10; Gaps 2;

Qy 5 PRPGAPCRAGA----PTIVLTSGRRQTLSHGSSSPARATLGKP 43

Db 76 PRPTPCSGSRWRPPT-----RRPTPPSTSPPCRSGTP 112

RESULT 44
US-10-156-761-10072
; Sequence 10072, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10072
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10072

Query Match 22.8%; Score 61; DB 4; Length 595;
Best Local Similarity 34.8%; Pred. No. 1.5e+02;
Matches 16; Conservative 3; Mismatches 27; Indels 0; Gaps 0;

Qy 2 AAGPRPGAPCRAGATTIVLTSGRRQTLTSHGSSSPARATLGKPLVLN 47
Db 487 AANGRPQGRGVGAPEPTARFGSSATASRGAGTSEAVTGRPSARN 532

RESULT 45
US-10-108-260A-2611
; Sequence 2611, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2611
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2611

Query Match 22.8%; Score 61; DB 4; Length 656;
Best Local Similarity 36.4%; Pred. No. 1.7e+02;
Matches 20; Conservative 6; Mismatches 21; Indels 8; Gaps 3;

Qy 2 AAGPRPGAPCRAGATTIVLTSGRRQTLTSHG-----SSSPARATLGKPLVLNDYS 50
Db 348 ASSPSPAP-RPERPESILVSGPSVTLTEGLGTVRPEQDPAKSP-GSPFLLRGLS 400

Search completed: December 27, 2005, 23:18:33
Job time : 165 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 23:05:09 ; Search time 12 Seconds
(without alignments)
30.909 Million cell updates/sec

Title: US-09-819-144A-2

Perfect score: 268

Sequence: 1 MAAGPRGAPCRAGAPTIVL.....SSPARATLGKPLVLDYSLN 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 54001 seqs, 7132810 residues

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New.*
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.5	21.5	1614	7	US-11-108-528-82
2	57	21.3	575	6	US-10-131-826A-128
3	56	20.9	1306	6	US-10-995-561-1027
4	54.5	20.3	368	7	US-11-000-463-916
5	54.5	20.3	388	7	US-11-000-463-444
6	54	20.1	138	6	US-10-667-295-176
7	54	20.1	2011	7	US-11-080-991-56
8	53	19.8	1166	6	US-11-821-234-964
9	53	19.8	1466	7	US-11-186-284-33
10	51.5	19.2	184	6	US-10-467-657-2790
11	51.5	19.2	259	6	US-10-512-184-34
12	51.5	19.2	269	7	US-11-015-546A-10
13	51.5	19.2	371	6	US-10-512-184-71
14	51.5	19.2	626	6	US-10-512-184-49
15	51.5	19.2	1493	7	US-11-004-057-4
16	51.5	19.2	1493	7	US-11-004-057-21
17	51.5	19.2	3690	6	US-10-995-561-1016
18	51.5	19.2	3714	6	US-10-995-561-1015
19	51.5	19.2	3717	6	US-10-821-234-1076
20	51	19.0	218	7	US-11-143-980-33
21	51	19.0	1618	6	US-10-984-645-2
22	50.5	18.8	637	6	US-10-821-234-961
23	50	18.7	529	6	US-10-821-234-1168
24	50	18.7	594	6	US-10-131-826A-10
25	50	18.7	2630	7	US-11-186-731-2

26	50	18.7	7968	7	US-11-186-731-5
27	49.5	18.5	177	6	US-10-999-866-61
28	49.5	18.5	205	6	US-10-995-561-1028
29	49.5	18.5	205	6	US-10-995-561-1029
30	49.5	18.5	367	6	US-10-131-826A-208
31	49.5	18.5	441	6	US-10-467-657-3438
32	49.5	18.5	480	7	US-11-132-142-6
33	49.5	18.5	1464	7	US-11-000-463-243
34	49.5	18.5	1464	7	US-11-186-284-28
35	49.5	18.5	1467	6	US-10-821-234-1096
36	49.5	18.5	1823	6	US-10-995-561-988
37	49.5	18.5	2102	6	US-10-995-561-990
38	49.5	18.5	2108	6	US-10-995-561-989
39	49.5	18.5	2157	6	US-10-995-561-991
40	49.5	18.5	3063	7	US-11-186-284-26
41	49	18.3	407	6	US-10-995-561-1014
42	49	18.3	707	7	US-11-186-284-132
43	48.5	18.1	188	6	US-10-467-657-920
44	48	17.9	432	6	US-10-467-657-7234
45	48	17.9	599	7	US-11-109-157A-3

ALIGNMENTS

RESULT 1
US-11-108-528-82
; Sequence 82, Application US/11108528
; Publication No. US20050261189A1
; GENERAL INFORMATION:
; APPLICANT: Larsen, Glenn
; APPLICANT: Marvin, Martha
; APPLICANT: Li, Dean Y.
; APPLICANT: Wang, Elizabeth
; APPLICANT: Chen, C. M. Amy
; APPLICANT: Shamah, Steven M.
; TITLE OF INVENTION: METHODS OF PROMOTING CARDIAC CELL
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: HYDR-P01-041
; CURRENT APPLICATION NUMBER: US/11/108,528
; PRIOR FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: US 60/563,137
; PRIOR FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/598,368
; PRIOR FILING DATE: 2004-08-02
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 1614
; TYPE: PRT
; ORGANISM: Mouse
US-11-108-528-82

Query Match 21.5%; Score 57.5; DB 7; Length 1614;
Best Local Similarity 32.1%; Pred. No. 16;
Matches 17; Conservative 7; Mismatches 18; Indels 11; Gaps 2;

QY 4 GPRGAPC-RAGAPTIVLTSGRRQT-----LSHGSSSPARATLGKPLV 45
DB 1445 GPRGAPC-RAGAPTIVLTSGRRQT-----LSHGSSSPARATLGKPLV 1497

RESULT 2
US-10-131-826A-128
; Sequence 128, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang

```
; APPLICANT: Gerriteen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 128
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-128

Query Match      21.3%; Score 57; DB 6; Length 575;
Best Local Similarity 31.5%; Pred. No. 5.7;
Matches 17; Conservative 7; Mismatches 18; Indels 12; Gaps 2;

Qy      1  MAAGPRP-----GAPCRAGAPTIVLTSGRRQTLTSHGSSSPARATLGGK 43
Db      113 LSTGPAAPAAATTSSKPEGRGQAAPTILLTKPGAT-SRPTTAPRTTTRP 165

RESULT 3
US-10-995-561-1027
; Sequence 1027, Application US/109955561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1027
; LENGTH: 1306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-1027
```

```
Query Match      20.9%; Score 56; DB 6; Length 1306;
Best Local Similarity 32.0%; Pred. No. 20;
Matches 16; Conservative 10; Mismatches 22; Indels 2; Gaps 1;

Qy      3  AGPRPGAPCRAGAPTIVLTSGRRQTLTSHGSSSPARATLGGKPLVNDYSLN 52
Db      17  AGLGSGAPALSG--SCLLTSGRCRCHRRHTALSHHPPLARPLFWSQLFLH 64

RESULT 4
US-11-000-463-916
; Sequence 916, Application US/110000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 916
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-916

Query Match      20.3%; Score 54.5; DB 7; Length 368;
Best Local Similarity 52.0%; Pred. No. 7;
Matches 13; Conservative 1; Mismatches 8; Indels 3; Gaps 1;

Qy      4  GPRPGAPCRAGAPTIVLT---SGRR 25
Db      254  GPRPGHPKGAFAQLNRSLSGQR 278

RESULT 5
US-11-000-463-444
; Sequence 444, Application US/110000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
```



```
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 1466
; TYPE: PRT
; ORGANISM: Homo Sapiens
;
US-11-186-284-33

Query Match          19.8%; Score 53; DB 7; Length 1466;
Best Local Similarity 42.9%; Pred. No. 53;
Matches 18; Conservative 3; Mismatches 17; Indels 4; Gaps 3;

Qy 2 AAGPRPGACRAGAPTIVLTSGRRQTLSHGSSSPARATLGRP 43
Db 709 AAGP-PGPPGAAGTGPLQMPGERGGL--GSPGP-KGDKGRP 746

RESULT 10
US-10-467-657-2790
; Sequence 2790, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2790
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
;
US-10-467-657-2790

Query Match          19.2%; Score 51.5; DB 6; Length 184;
Best Local Similarity 28.9%; Pred. No. 7.5;
Matches 13; Conservative 7; Mismatches 20; Indels 5; Gaps 1;

Qy 13 AGAPTIVLTSGRRQTLSS-----HGSSSPARATLKGPLVLDYSLN 52
Db 49 SAADNSIMTKQKYVESNCIACHKGKGGRGTAFPPFLFRSDYIMN 93

RESULT 11
US-10-512-184-34
; Sequence 34, Application US/10512184
; Publication No. US20050244901A1
; GENERAL INFORMATION:
; APPLICANT: Fraunhofer Gesellschaft zur Forderung der angewandten Forschung e.V.
; TITLE OF INVENTION: Antibodies, recombinant antibodies, recombinant
; TITLE OF INVENTION: antibody fragments and fusions mediated plant disease
; FILE REFERENCE: 3581.01US01
; CURRENT APPLICATION NUMBER: US/10/512,184
; CURRENT FILING DATE: 2004-10-22
; SOFTWARE: PatentIn Ver. 2.1
```

```
; TITLE OF INVENTION: resistance against fungi
; FILE REFERENCE: 3581.01US01
; CURRENT APPLICATION NUMBER: US/10/512,184
; CURRENT FILING DATE: 2004-10-22
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: scFv PL2 with
; OTHER INFORMATION: specificity against Phoma lingam; originates from
; OTHER INFORMATION: Mus musculus.
;
US-10-512-184-34

Query Match          19.2%; Score 51.5; DB 6; Length 259;
Best Local Similarity 39.5%; Pred. No. 11;
Matches 15; Conservative 3; Mismatches 11; Indels 9; Gaps 1;

Qy 5 PRPGACRAGAPTIVLTSS-----GRRQTLSHGSS 33
Db 133 PGFGSGTKGAPRIVLTQSPSSSLAMSVGQKVTMSCKSS 170

RESULT 12
US-11-015-546A-10
; Sequence 10, Application US/11015546A
; Publication No. US20050250126A1
; GENERAL INFORMATION:
; APPLICANT: GAO, ZEREN
; APPLICANT: SHEPPARD, PAUL O.
; APPLICANT: FOX, BRIAN A.
; APPLICANT: HOLLOWAY, JAMES L.
; APPLICANT: JASPERS, STEPHEN R.
; APPLICANT: APPELEY, MARK
; TITLE OF INVENTION: Ztnf13, A TUMOR NECROSIS FACTOR
; FILE REFERENCE: 03-24
; CURRENT APPLICATION NUMBER: US/11/015,546A
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 60/530,185
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Mus musculus
;
US-11-015-546A-10

Query Match          19.2%; Score 51.5; DB 7; Length 269;
Best Local Similarity 27.7%; Pred. No. 12;
Matches 13; Conservative 6; Mismatches 9; Indels 19; Gaps 2;

Qy 1 MAAGPRPGACRAGAPTIVLTSGR-----RQTLSHGS 32
Db 211 LPSAPRPAAPC-----PALCLQAERSRRVFPSTDSLSRPLEHGA 253

RESULT 13
US-10-512-184-71
; Sequence 71, Application US/10512184
; Publication No. US20050244901A1
; GENERAL INFORMATION:
; APPLICANT: Fraunhofer Gesellschaft zur Forderung der angewandten Forschung e.V.
; TITLE OF INVENTION: Antibodies, recombinant antibodies, recombinant
; TITLE OF INVENTION: antibody fragments and fusions mediated plant disease
; FILE REFERENCE: 3581.01US01
; CURRENT APPLICATION NUMBER: US/10/512,184
; CURRENT FILING DATE: 2004-10-22
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
```



```
US-10-995-561-1015
; Sequence 1015, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1015
; LENGTH: 3714
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-1015

Query Match      19.2%; Score 51.5; DB 6; Length 3714;
Best Local Similarity 31.1%; Pred. No. 2.4e+02;
Matches 19; Conservative 5; Mismatches 20; Indels 17; Gaps 3;

QY      3 AGPRPGAP-----CRAGPTIVLTS-----GRRQTLSHGSSSPARATLGKPLVL 46
Db      1362 AGPRPRPTAVDPDAEPTLLREPOATVFTTHVPTLGRYAFLLHG-YQPAHPTFPVEVLI 1320
QY      47 N 47
Db      1321 N 1321

RESULT 19
US-10-821-234-1076
; Sequence 1076, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 1076
; LENGTH: 3717
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1076

Query Match      19.2%; Score 51.5; DB 6; Length 3717;
Best Local Similarity 31.1%; Pred. No. 2.4e+02;
Matches 19; Conservative 5; Mismatches 20; Indels 17; Gaps 3;

QY      3 AGPRPGAP-----CRAGPTIVLTS-----GRRQTLSHGSSSPARATLGKPLVL 46
Db      1283 AGPRPRPTAVDPDAEPTLLREPOATVFTTHVPTLGRYAFLLHG-YQPAHPTFPVEVLI 1341
QY      47 N 47
Db      1342 N 1342

RESULT 20
US-11-143-980-33
; Sequence 33, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
```

```
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haltli, Bradley A.
; APPLICANT: Wagenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; TITLE OF INVENTION: Polyketide
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; PRIOR FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Streptomyces sp.
US-11-143-980-33

Query Match      19.0%; Score 51; DB 7; Length 218;
Best Local Similarity 44.0%; Pred. No. 11;
Matches 11; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY      1 MAAGPRPGAPCRAGAPT--IVLTSG 23
Db      1 VASPPSPASPARPGRPVRLVLSG 25

RESULT 21
US-10-984-645-2
; Sequence 2, Application US/10984645
; Publication No. US20050244386A1
; GENERAL INFORMATION:
; APPLICANT: Habener, Joel
; APPLICANT: Zulewski, Hendrik
; APPLICANT: Abraham, Elizabeth
; APPLICANT: Vallejo, Mario
; TITLE OF INVENTION: METHOD OF TRANSPLANTING IN A MAMMAL AND TREATING DIABETES MELLITUS
; TITLE OF INVENTION: BY ADMINISTERING A PSEUDO-ISLET LIKE AGGREGATE DIFFERENTIATED FROM
; TITLE OF INVENTION: A NESIN-POSITIVE PANCREATIC STEM CELL
; FILE REFERENCE: 3284/1223
; CURRENT APPLICATION NUMBER: US/10/984,645
; CURRENT FILING DATE: 2004-11-09
; PRIOR APPLICATION NUMBER: US 09/731,255
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 60/169,082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215,109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/239,880
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-984-645-2

Query Match      19.0%; Score 51; DB 6; Length 1618;
Best Local Similarity 35.8%; Pred. No. 1.1e+02;
Matches 19; Conservative 3; Mismatches 25; Indels 6; Gaps 2;

QY      5 PRP-GAPCRAGATIVL-----TSGRRQTLSHGSSSPARATLGKPLVINDYSL 51
Db      424 PEPLRAEARVAIPASVLPGPEFGGQRQEAATGQSPEDHASLAPPLSPDHSSL 476
```


Thu Dec 29 19:53:30 2005

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowsky, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330RIC128
; CURRENT APPLICATION NUMBER: US/10/131.826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-131-826A-10

Query Match 18.7%; Score 50; DB 6; Length 594;
Best Local Similarity 37.8%; Pred. No. 45;
Matches 14; Conservative 6; Mismatches 13; Indels 4; Gaps 1;

QY 3 AGPRGAPCRAGAPTIVLTSGRRQTLGSHGSSSPARATL 39
Db 358 APPQPGP-----APGRNASARRRRSSSSSSSSASRTS 390

RESULT 25
US-11-186-731-2
; Sequence 2, Application US/11186731
; Publication No. US2005025521A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047PIRCPI(M)
; CURRENT APPLICATION NUMBER: US/11/186,731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowsky, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330RIC128
; CURRENT APPLICATION NUMBER: US/10/131.826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-131-826A-10

Query Match 18.7%; Score 50; DB 6; Length 529;
Best Local Similarity 29.2%; Pred. No. 39;
Matches 21; Conservative 6; Mismatches 25; Indels 20; Gaps 2;

QY 1 MAAGPR-PGAPCRAGAPTIVLTSGRRQTL-----SHGSSSPARATL 40
Db 5 LLAGPRLPGVPGAGQAQSMAGAGPKRRALAAPAAEEKEERKWLAAKADGSAPAGEGE 64

QY 41 GKPLVLDYSLN 52
Db 65 GVTLQRNITLLN 76

RESULT 24
US-10-131-826A-10
; Sequence 10, Application US/10131826A
; Publication No. US20050245730A1
```

```
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-731-2

Query Match      18.7%; Score 50; DB 7; Length 2630;
Best Local Similarity 32.7%; Pred. No. 2.5e+02;
Matches 16; Conservative 1; Mismatches 18; Indels 14; Gaps 2;

Qy  2 AAGPRPG-----AFCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLV 45
Db  905 AFGPSTGDLTGPGCPRGAPALQET-----GSQPPVTGTSEAPAV 944

RESULT 26
US-11-186-731-5
; Sequence 5, Application US/11186731
; Publication No. US2005025521A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 53079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047PIRCPl(M)
; CURRENT APPLICATION NUMBER: US/11/186,731
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 7968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-731-5

Query Match      18.7%; Score 50; DB 7; Length 7968;
Best Local Similarity 32.7%; Pred. No. 8.8e+02;
Matches 16; Conservative 1; Mismatches 18; Indels 14; Gaps 2;

Qy  2 AAGPRPG-----AFCRAGAPTIVLTSGRRQTLSHGSSSPARATLGKPLV 45
Db  6243 AFGPSTGDLTGPGCPRGAPALQET-----GSQPPVTGTSEAPAV 6282

RESULT 27
US-10-999-866-61
; Sequence 61, Application US/10999866
; Publication No. US20050266004A1
; GENERAL INFORMATION:
; APPLICANT: GILES-KOMAR, Jill; SCALLON, Bernard J.; CAI, Ann
; TITLE OF INVENTION: ANTI-HUMAN LYMPHOTOXIN ALPHA ANTIBODIES, COMPOSITIONS, METHODS AN
; FILE REFERENCE: CENS042NP
; CURRENT APPLICATION NUMBER: US/10/999,866
; CURRENT FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: 60/527,794
; PRIOR FILING DATE: 2003-12-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
```

```
; LOCATION: (1)..(177)
; OTHER INFORMATION: LTA
US-10-999-866-61

Query Match      18.5%; Score 49.5; DB 6; Length 177;
Best Local Similarity 33.3%; Pred. No. 13;
Matches 15; Conservative 4; Mismatches 15; Indels 11; Gaps 2;

Qy  7 PGAPCRAGAPTIVLTSGRRQT-----LSHGSSSPARATLGKP 43
Db  2 PGA---QGLPGVGLTPSAAQTARQHPKMHLSHTLKPAAHLIGDP 43

RESULT 28
US-10-995-561-1028
; Sequence 1028, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1028
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-1028

Query Match      18.5%; Score 49.5; DB 6; Length 205;
Best Local Similarity 33.3%; Pred. No. 15;
Matches 15; Conservative 4; Mismatches 15; Indels 11; Gaps 2;

Qy  7 PGAPCRAGAPTIVLTSGRRQT-----LSHGSSSPARATLGKP 43
Db  30 PGA---QGLPGVGLTPSAAQTARQHPKMHLSHTLKPAAHLIGDP 71

RESULT 29
US-10-995-561-1029
; Sequence 1029, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1029
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-1029

Query Match      18.5%; Score 49.5; DB 6; Length 205;
Best Local Similarity 33.3%; Pred. No. 15;
Matches 15; Conservative 4; Mismatches 15; Indels 11; Gaps 2;

Qy  7 PGAPCRAGAPTIVLTSGRRQT-----LSHGSSSPARATLGKP 43
Db  30 PGA---QGLPGVGLTPSAAQTARQHPKMHLSHTLKPAAHLIGDP 71

RESULT 30
US-10-131-826A-208
```

```
; Sequence 208, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 208.
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-208

Query Match      18.5%; Score 49.5; DB 6; Length 367;
Best Local Similarity 31.8%; Pred. No. 30;
Matches 14; Conservative 4; Mismatches 25; Indels 1; Gaps 1;

QY 3 AGPRGACACAGATTIVLTSGRRQTLTSHGSSSPARATLTKPLVL 46
Db 76 SGTTFSEPSAPAPAAVAPRLSGSNH-SGSPKLTGKRLPOL 118

RESULT 31
US-10-467-657-3438
; Sequence 3438, Application US/10467657
; Publication No. US20050260591A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3438
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3438

Query Match      18.5%; Score 49.5; DB 6; Length 441;
Best Local Similarity 34.0%; Pred. No. 37;
Matches 18; Conservative 4; Mismatches 16; Indels 15; Gaps 2;

QY 2 AAGPRGACACAGAPT-----IVLTSGRRQTLTSHGSSS-----PARAT 39
Db 74 AASPAPRPPTVHAAPTQTPTETLTAHVLDLSRQTLTSAKNTTTPVEPALT 126

RESULT 32
US-11-132-142-6
; Sequence 6, Application US/11132142
; Publication No. US20050276818A1
; GENERAL INFORMATION:
; APPLICANT: The Burnham Institute
; APPLICANT: Sikora, Sergey
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: UNCHARACTERIZED ORF3 IN SARS-CORONAVIRUS IS A
; TITLE OF INVENTION: CYCLIC-AMP-DEPENDENT KINASE AND A TARGET FOR SARS THERAPY
; FILE REFERENCE: 8014-011-US
; CURRENT APPLICATION NUMBER: US/11/132,142
; CURRENT FILING DATE: 2005-05-17
; PRIOR APPLICATION NUMBER: 60/571,698
; PRIOR FILING DATE: 2004-05-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 6
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Aspergillus sp.
US-11-132-142-6

Query Match      18.5%; Score 49.5; DB 7; Length 480;
Best Local Similarity 35.1%; Pred. No. 40;
Matches 20; Conservative 6; Mismatches 18; Indels 13; Gaps 4;

QY 5 PRPGA-----PCRAGAPTIVLTSGRRQTLTSHGSSSPARATLTKPLVINDYSL 51
Db 118 PQPNANAASIQNIINPSQOGAMHSA-SSGHTQS-HHAGRSDARTTKGR-YSLDDFS 171

RESULT 33
US-11-000-463-243
; Sequence 243, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
```



```
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 2102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-990

Query Match      18.5%; Score 49.5; DB 6; Length 2102;
Best Local Similarity 39.5%; Pred. No. 2.2e+02;
Matches 15; Conservative 3; Mismatches 15; Indels 5; Gaps 2;

Qy 5 PRPGAPCRAGAPTIVLTSGRRQTLSHGS--SSPARATL 40
Db 504 PTPGV---AAAPTIVSGGSTSTSSGSGFEASFPVEPQL 538

RESULT 38
US-10-995-561-989
; Sequence 989, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 989
; LENGTH: 2108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-989

Query Match      18.5%; Score 49.5; DB 6; Length 2108;
Best Local Similarity 39.5%; Pred. No. 2.2e+02;
Matches 15; Conservative 3; Mismatches 15; Indels 5; Gaps 2;

Qy 5 PRPGAPCRAGAPTIVLTSGRRQTLSHGS--SSPARATL 40
Db 504 PTPGV---AAAPTIVSGGSTSTSSGSGFEASFPVEPQL 538

RESULT 39
US-10-995-561-991
; Sequence 991, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 991
; LENGTH: 2157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-991

Query Match      18.5%; Score 49.5; DB 6; Length 2157;
Best Local Similarity 39.5%; Pred. No. 2.3e+02;
Matches 15; Conservative 3; Mismatches 15; Indels 5; Gaps 2;

Qy 5 PRPGAPCRAGAPTIVLTSGRRQTLSHGS--SSPARATL 40
Db 504 PTPGV---AAAPTIVSGGSTSTSSGSGFEASFPVEPQL 538

RESULT 40
US-11-186-284-26
; Sequence 26, Application US/11186284
; Publication No. US20050286493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 3063
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-186-284-26

Query Match      18.5%; Score 49.5; DB 7; Length 3063;
Best Local Similarity 36.8%; Pred. No. 3.4e+02;
Matches 14; Conservative 5; Mismatches 16; Indels 3; Gaps 2;

Qy 2 AAGPR--PGAPCRAGAPTIVLTSGR-RQTLSHGSSSPA 36
Db 2849 AMGPRGPRGPGSPGVTGPGKPGKPGDHRPGPS 2886

RESULT 41
US-10-995-561-1014
; Sequence 1014, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1014
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-1014

Query Match      18.3%; Score 49; DB 6; Length 407;
Best Local Similarity 35.4%; Pred. No. 39;
Matches 17; Conservative 8; Mismatches 13; Indels 10; Gaps 3;
```


This Page Blank (usps10)